ESSAYS ON MACRO-FINANCE

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

CHAO YING

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

ROBERT GOLDSTEIN

May, 2021

© CHAO YING 2021
ALL RIGHTS RESERVED

Acknowledgements

I am deeply indebted to many others who helped me throughout graduate school.

I cannot possibly say enough about Bob's role in my growth as a researcher and the work underlying this thesis. Without his above-and-beyond support, I literally would not have been able to graduate. I have learned an immense amount from working with him over the past several years, and I appreciate that he treats me as an equal partner during our work.

I feel so regret that I did not work with Colin until my 4th year. It is a pleasant journey working with people who share the same research interests and, more importantly, views of life and value. I may forget the shape of PDE boundaries in the future but definitely not the log cabin he built and painted for his children.

Xiaoji has enriched all of us in Minnesota by returning here. His level of commitment to students is far beyond most faculty, and it will be an inspiration to me as I set out along the same path. I want to especially express my appreciation for his support in the face of the adversities of my life.

I would like to thank Erzo for encouraging me to pursue a Ph.D. and many helpful suggestions and extensive discussions. Besides, I really appreciate the help and encouragement from my wonderful coauthors Luca and Lorenzo; their words of advice and comfort at the time were just as precious as those of my committee members.

I want to thank Erik, Martin, and Tracy for their help during my PhD study and the job market. I have had a lot of fun talking with Jash, Ramin, and Yu. I have also met many nice scholars at the conferences and during the job market. Special thanks go to my hero Kerry, my brilliant coauthors Yunzhi and Felipe, and my knowledgeable discussants Christian and Julio.

I am thankful for Batu, Han, and Ran for the most memorable times in Minnesota.

I am so lucky to have Qinfang and Shengliang, who always support and encourage me no matter when.

Finally, I would like to thank my family, especially my parents and my significant other, for their unwavering love and support.

Dedication

To those who held me up over the years

Abstract

My dissertation investigates the interaction between macroeconomics and finance. It contains three chapters.

Chapter 1 studies the private information explanation for the time-series pre-FOMC drift through risk reduction. Using transaction-level data, I document the informed trading is in the same direction of the realized returns in the 24-hour window before FOMC announcements, coinciding with the pre-FOMC uncertainty reduction. I integrate Kyle's (1985) model into a standard consumption-based asset pricing framework where the market makers are compensated for the risk of assets' fundamentals. Observing aggregate order flow, they update the belief about the marginal utility-weighted asset value, which resolves uncertainty gradually and results in an upward drift in market prices before announcements. I demonstrate that there is a strictly positive pre-FOMC drift if and only if the market makers require risk compensation.

Chapter 2 is co-authored with Colin Ward. We develop an equilibrium model where cash holdings, costly refinancing policies, and managerial incentives are jointly determined to quantify the market's influence on management's ex ante behavior. We also derive a general formula that shows how agency and financing distortions shape payouts and compensation, two easily measured quantities. Our calibrated model estimates agency conflicts are nearly 10 times more severe than financial frictions for US public firms. Our analysis suggests that cutting corporate income taxes while introducing a tax on refinancing can reduce the relative severity of agency.

Chapter 3 is joint work with Luca Benzoni, Lorenzo Garlappi, Robert S. Goldstein, and Julien Hugonnier. We investigate the optimal dynamic debt policy of a firm that issues non-callable debt subject to a fixed cost when shareholders cannot commit to future restructuring policies. We derive necessary and sufficient conditions for the existence of no-commitment Markov perfect equilibria. For a given debt issuance cost parameter, we identify a range of maturities for which equilibria exists and tax benefits are positive. In particular, for realistic values of issuance costs and debt maturity, our no-commitment framework generates equity and debt prices that are only slightly lower than those obtained under a global-optimal policy with commitment.

Contents

List of Tables

List of Figures

Chapter 1

The Pre-FOMC Announcement Drift and Private Information: Kyle Meets Macro-Finance

Lucca and Moench (2015) document the substantial stock market returns before the Federal Open Market Committee (FOMC) announcements. They find that the pre-FOMC drift of the S&P 500 index is on average 49 basis points during the 24-hour window preceding FOMC announcements, which corresponds to about 80% of the annual realized excess returns in the stock market. However, the hours and days before FOMC meetings fall into the "blackout period," a time when policymakers and Fed staff refrain from discussions of monetary policy information.[1] It provides a notable challenge to standard asset pricing theory, which predicts equity returns should be earned *at* the announcements when uncertainty is resolved from the public news, *rather than ahead*.

Some papers have shown suggestive evidence that the pre-FOMC drift may come from private information before announcements. Cieslak, Morse, and Vissing-Jorgensen (2019) and Vissing-Jorgensen (2020) provide a history of leak discussions in FOMC documents and argue that systematic information leakage drives the pre-FOMC announcement drift. In addition to information leakage, market participants may generate their

[1] The blackout period begins at the start of the second Saturday (midnight) Eastern Time before the beginning of the meeting and ends at midnight Eastern Time on the next day after the meeting.

proprietary information by collecting data related to FOMC announcements.[2] In this paper, I study the private information explanation for the time-series pre-FOMC drift through risk reduction.

Empirically, I provide asset-market-based evidence that supports the presence of private information and insider trading before FOMC announcements. First, Hu, Pan, Wang, and Zhu (2020) find the significant and systematic reduction of market uncertainty (measured by the CBOE VIX index) during the same 24-hour window before FOMC announcements. Second, sorting the FOMC days via 24-hour uncertainty reduction before announcements into terciles, I find that only the group with substantial uncertainty reduction preceding announcements are associated with the positive pre-FOMC drift. Third, to measure informed trading, I calculate the order imbalances, defined as the difference between buyer- and seller- initiated trading volumes divided by total trading volume. When the uncertainty reduces before FOMC, the abnormal order imbalances are 1.85-2.17% higher in the direction of the realized return in the 24-hour window before FOMC announcements.

To understand the above features of the financial markets, I build a model that the pre-FOMC announcement drift is earned as risk reduces through insider trading. I integrate Kyle's (1985) model into a standard consumption-based asset pricing framework such that the market makers are compensated for the risk of assets' fundamentals. The equity premium is realized with uncertainty reduction prior to announcements since insider trading reveals private information. Characterizing the equilibrium price and insider trading by closed-form, I establish a strictly positive pre-FOMC announcement drift if and only if the market makers are *risk-compensated.*

Since FOMC announcements provide information about the macro-economy, the market makers require risk compensation in assets' fundamentals before the announcements. To provide macroeconomic conditions for the market makers' pricing decisions, I develop a continuous-time equilibrium model in which the aggregate economic growth is driven by a latent state variable and an i.i.d. component (short-run shocks). The investors cannot observe the latent variable directly and update the belief by observing the aggregate endowment when there is no announcement. The competitive market makers

[2] For example, Kurov, Sancetta, Strasser, and Wolfe (2017) show that proprietary information permits forecasting announcement surprises in some cases.

set the price, which equals the marginal utility weighted payoffs through the stochastic discount factor (SDF) determined from the above economy. The counter-cyclical SDF applies extra discounting to payoffs positively correlated with utility. Thus, the asset market requires a premium for such payoffs relative to risk-free returns.

The Fed has some extra knowledge regarding the economy, which should be revealed through periodic FOMC announcements. However, the insider knows the underlying information before announcements and trades to maximize the expected terminal profit, understanding the order affects the price. Meanwhile, the liquidity traders have random, price-inelastic demands as in the standard Kyle model. By observing aggregate order flow, the market markers update the estimation of asset payoffs as well as the SDF simultaneously such that their uncertainty is resolved.

Here are some implications of the equilibrium with the risk-compensated market makers. First, the equilibrium price is a submartingale, instead of a martingale in the standard continuous-time Kyle-type models.[3] The intuition is as follows. Because of risk compensation, on average, the price of risky assets increases as uncertainty is resolved through insider trading before announcements. The slope of the expected pre-FOMC drift is the negative variance between the innovation of the SDF and the asset value. I prove a strictly positive pre-FOMC drift if and only if the market makers are compensated for the risk of assets' fundamentals. The positive excess return leads to positive average order imbalances before announcements. In the meantime, to entice the insider to trade and release information early, the market makers have incentives to set the price impact that increases on average, implying the submartingale property of the price impact. Second, due to the average upward drift in market prices, the market makers rationally anticipate that the insider would trade positively on average to chase that premium. The insider also has to consider the additional price impact from uncertainty resolution via her trading, which is unique in this model. In the equilibrium, instead of being zero, the expected order rate is determined by the ratio of the pre-FOMC drift's slope to Kyle's lambda. Additionally, as the market makers converge to be risk-neutral, the limit of the equilibrium is well defined and converges to the traditional Kyle model. The equilibrium implications indicate that this paper provides a microfoundation of how private news diffuses drive positive pre-FOMC drift

[3] See Back (1992), Back and Pedersen (1998), Li (2013), and Collin-Dufresne and Fos (2016), etc.

in a standard microstructure framework.

Since uncertainty is not always reduced before FOMC meetings, I generalize the benchmark model so that the insider may not be better informed, and the market makers assess whether the insider has private information or not. In addition to updating their belief of asset payoffs and the SDF, the market makers estimate the probability that insiders have private information simultaneously, which are solved by a nonlinear filtering technique. Conditional on the insider being informed, the closed-form equilibrium price is a submartingale as the benchmark when the maker makers are risk-compensated. The growth rate of the expected pre-FOMC announcement drift and the expected insider's order rate are time-varying, caused by the dynamics of the probability estimate. The pricing rule is nonlinear and stochastic, which drives price volatility, market depth, and price response to be stochastic. I calibrate the model so that stochastic pricing dynamics are consistent with both the *level* and the *trend* of the time-varying 24-hour pre-FOMC announcement drift in the data.

Before concluding, I demonstrate that other asset market evidence around FOMC announcements is consistent with the model's predictions. First, empirically the pre-FOMC drift is stronger when the uncertainty reduces more before announcements, coinciding with the risk-based explanation. Second, to maximize her profits, the insider faces a tradeoff between uncertainty and liquidity. She wants to trade later such that the market uncertainty is higher. However, she can not trade too late since she needs substantial liquidity trading to hide her position. It explains the pre-FOMC drift's timing that occurs 24 hours before announcements when private information is probably known way before. Third, I document the market uncertainty decreases significantly only before announcements with press conferences since April 2011, which explains the two distinctive patterns to equity returns found in Boguth, Gregoire, and Martineau (2019) and agrees with my model. Fourth, the sign of the pre-FOMC drift in the model depends on whether the asset is risky or a hedge. The average of time-varying betas of nominal bond is close to zero from 1996 to 2019, resulting in the absence of the pre-FOMC drift in fixed income instruments.

Related literature

The paper relates to several strands of the literature. First is the large body of work investigating the impact of asymmetric information on asset prices and price impacts, seminal examples of which include Kyle (1985) and Back (1992).[4] I build on this literature by exploring the implications of the risk-compensated market makers, which have been largely ignored in the literature.[5] The equilibrium price in this model is a submartingale instead of a martingale since the resolution of uncertainty is associated with the realizations of the premium. Meanwhile, the market makers rationally anticipate that the insider would trade positively on average to chase that premium. The insider also has to consider this additional price impact from uncertainty resolution when she trades, which is unique in my model. The dynamic game between the market makers and the insider results in a positive expected order rate from the insider that contrasts this model to the literature. The limit of the equilibrium is the traditional Kyle model as the market makers converge to be risk-neutral, establishing the link between this paper and the literature.[6]

This paper exploits the main insight in macro-finance literature, which addresses the importance of macroeconomic conditions to account for asset prices. Starting from the equity premium puzzle in Mehra and Prescott (1985) and Hansen and Jagannathan (1991), the literature has explored a wide range of alternative preferences and market structures to account for the equity market dynamics.[7] However, very little attention has been paid to the pre-FOMC drift, which corresponds to about 80% of the annual realized excess returns in the stock market.[8] Integrating Kyle's (1985) model into

[4] A short list is Foster and Viswanathan (1996), Back and Pedersen (1998), Back, Cao, and Willard (2000), Caldentey and Stacchetti (2010), Li (2013), Collin-Dufresne and Fos (2016), Back, Crotty, and Li (2018), Drechsler, Moreira, and Savov (2018), Dai, Wang, and Yang (2019), Crego (2020).

[5] While Subrahmanyam (1991) considers risk-averse market makers under CARA utility in a one-period Kyle model, there is no pre-announcement drift since the fundamental risk in the news is not priced due to CARA utility.

[6] This paper is also connected to inventory models, such as Stoll (1978) and Ho and Stoll (1981). They consider a risk-averse market maker that holds undesired positions and requires the compensation in terms of a bid-ask spread. However, these papers abstract from information asymmetries. Vayanos (2001) studies a model with risk-averse market makers that the large trader's private information is about her endowment shocks rather than the fundamental value of asset.

[7] See reviews in Cochrane (2017).

[8] Section 1.6 talks about the details of other explanations, including Hu, Pan, Wang, and Zhu (2020), Laarits (2019), and Cocoma (2020).

a standard consumption-based asset pricing framework, I establish a strictly positive pre-announcement drift if and only if the market makers are risk-compensated in the presence of insider trading. Therefore, this paper also fills the gap between macro-finance literature and microstructure literature related to Kyle (1985).

My paper contributes to the broader literature on the premium around FOMC announcements.[9] Savor and Wilson (2013) find a significant equity market return on days with major macroeconomic announcements.[10] Lucca and Moench (2015) document the substantial stock market return during the 24-hour period preceding FOMC announcements. Theoretically, Ai and Bansal (2018) provide a revealed preference theory for the macroeconomic announcement premium in a representative agent economy.[11] Given the market microstructure in Kyle's model, this paper accounts for both the level and the trend of the pre-FOMC drift in the presence of private information when the market makers are risk-compensated. The endogenous uncertainty reduction before announcements through insider trading agrees with the evidence documented in Hu, Pan, Wang, and Zhu (2019). This paper can be extended to study other pre-event drifts documented in the literature. A large group of papers treats average abnormal positive excess returns before events as evidence of insider trading and tests the market liquidity implications inspired by Kyle (1985), such as other macroeconomic announcements (Kurov, Wolfe, and Gilbert (2020)), mergers and acquisitions or earnings announcements (Keown and Pinkerton (1981), Penman (1982), and Meulbroek

[9] See more discussions in section 1.1.1, section 1.5, and section 1.6.

[10] Ernst, Gilbert, and Hrdlicka (2019) find the FOMC appears to stand out from the other macroeconomic announcements in all their results, which has the largest point estimates for the concentration of the equity premium. Giacoletti, Ramcharan, and Yu (2020) study the impact of FOMC announcements on the mortgage market.

[11] To account for the pre-FOMC announcement drift, Ai and Bansal (2018) assume the contents of announcements are communicated to the public a few hours before the pre-scheduled announcements, which leads all investors to receive informative signals before FOMC announcements. However, FOMC members refrain from discussions of monetary policy during this period, which implies that it is almost impossible that the public systemically receives information before announcements. Besides, the implication of this assumption is not consistent with other empirical facts upon FOMC announcements. For example, there are still monetary policy surprises (see Nakamura and Steinsson (2018)), huge trading volume as well as significant realized volatility (see Lucca and Moench (2015), Bollerslev, Li, and Xue (2018), and Ying (2020)) after FOMC announcements. None of these can happen if all investors know the information before announcements. Based on the generalized risk sensitivity in Ai and Bansal (2018), Ai, Bansal, Im, and Ying (2018) and Wachter and Zhu (2018) develop quantitative models of the announcement premium under a representative agent. Ying (2020) measures the impact of FOMC announcements on disagreement in a general equilibrium model with heterogeneous beliefs.

(1992)), and other events (Sinha and Gadarowski (2010), Agapova and Madura (2011), Collin-Dufresne and Fos (2015)). However, in the standard Kyle-type model, the expected average excess return before announcements is zero since the market makers are risk-neutral. Therefore, this paper provides a general theoretical framework for other pre-event drifts out of private information as long as the risk of the event is priced in the pricing kernel.

The rest of the paper is organized as follows. Section 1.1 provides empirical evidence that indicates the presence of private information before FOMC news. To provide macroeconomic conditions for the market makers' pricing decisions, I present the standard consumption-based asset pricing framework in section 1.2. Section 1.3 extends Kyle's (1985) model to the case where the market makers are risk-compensated and characterizes the equilibrium price and insider trading. In section 1.4, I generalize the benchmark model that the insider may not be informed, and the market makers assess whether the insider has private information or not. Section 1.5 tests the further implications of the model. I discuss the challenges of the private information explanation mentioned in the literature and talk about other explanations in section 1.6. Section 1.7 concludes. The Appendix contains additional details on the empirical analysis and the proof.

1.1 Empirical Evidence

In this section, I summarize the potential sources of private information before FOMC announcements and discuss the suggestive evidence shown in the literature. After that, I provide asset-market-based evidence that supports the presence of private information before FOMC news. First, I present that the market uncertainty (measured by the VIX index) decreases significantly and systematically during the same window as the pre-FOMC drift, as documented in Hu, Pan, Wang, and Zhu (2020). Second, sorting the FOMC days via 24-hour uncertainty reduction before announcements into terciles, I find that only the group with substantial uncertainty reduction preceding announcements are associated with the positive pre-FOMC drift. Third, I document there is only significant insider trading (measured by order imbalances) when uncertainty decreases before announcements, which is consistent with the private information explanation.

1.1.1 Sources of private information before FOMC meetings

The literature has provided suggestive evidence of private information before announcements. The private information may be obtained by leakage. Cieslak, Morse, and Vissing-Jorgensen (2019) propose that information about the Federal Reserve's unexpected accommodating monetary policy is leaked ahead of the FOMC announcement, which causes a pre-announcement equity market rally. Vissing-Jorgensen (2020) provides a history of leak discussions in FOMC documents to show that the FOMC itself expresses frequent concerns about leaks. For example, the leakage led to the resignation of Richmond Fed President Lacker following admission of her involvement in the leak of confidential FOMC information to Medley Global Advisers in 2012. Finer (2018) documents an abnormal number of NYC taxi rides to the district of liberty street certain times before FOMC announcements. Besides, the private information may come from the accidental information leakage—"word-of-mouth" interpretation of information diffusion, which has been well studied in the literature of takeovers (see Keown and Pinkerton (1981), Jarrell and Poulsen (1989), Meulbroek (1992), and Augustin, Brenner, and Subrahmanyam (2015)).

The other potential source of private information is through proprietary data collection related to FOMC announcements. Given the huge market attention to FOMC announcements, to infer what the Fed knows, institutional investors have strong motivations to obtain the information that the Fed observed and keep updating the prediction model of monetary policy from historical data.[12] Kurov, Sancetta, Strasser, and Wolfe (2017) support this explanation by finding that proprietary information permits forecasting announcement surprises in some cases.

1.1.2 The average cumulative VIX change and return before FOMC announcements

To capture the changes of market expectations in a timely manner, I use the CBOE VIX index, which is a model-free measure of implied volatility computed from the S&P 500 index option prices. For the intraday returns, I obtain transaction-level data on S&P

[12] For example, institutional investors can hire the talented, well-trained economists who help the Fed process and interpret all the information being released, as discussed in Nakamura and Steinsson (2018).

500 index (SPX). The sample period is from January 1996 to December 2019. During this period, there are in total 187 schuduled releases of FOMC statements. Except 9 of them, other releases are either around 2:15 p.m. ET (before April 2011) or 2:00 p.m. ET (after April 2011).[13] Therefore, I follow Lucca and Moench (2015) and focus on the 2 p.m.-to-2 p.m. pre-FOMC window, which should not contain any announcement information if there is no private information revealing.

Figure 1.1 shows the average cumulative VIX change and average cumulative return on the S&P 500 index around FOMC announcements. The solid line of the right panel represents the mean pointwise cumulative intraday percentage return of the SPX over a four-day window from the market open of the day ahead of scheduled FOMC meetings to the day after. Over the window from Day -3 through the beginning of Day -1, the average VIX increases due to the huge uncertainty of the upcoming FOMC news. While as shown in Table 1.2, the VIX decreases 0.3% with a t-stat of -3.4 during the 24-hour period preceding FOMC announcements, which is consistent with Hu, Pan, Wang, and Zhu (2020).[14] Meanwhile, the cumulative pre-FOMC drift over the same window is on average 33.2 basis points, which is statistically significant at the 1% level.

The significant reduction of VIX index shows the systemic uncertainty reduction out of the revealing of FOMC news preceding announcements. However, while it is common for FOMC members to express their views about macroeconomic developments or monetary policy issues in meetings or conversations with members of the public, they refrain from these discussions in the week before FOMC meetings. The 24-hour pre-FOMC window is part of the blackout period. Therefore, the significant uncertainty reduction in this window indicates the potential presence of private information before FOMC announcements.

[13] 8 of the 9 exceptions are released around 12:30 p.m. ET from April 2011 to December 2012. Another exception happened at 11:30 a.m. ET on March 26, 1996 because of the Chariman's other duties. The results hold robustly without these releases.

[14] The resolution of uncertainty occurs in two stages on the FOMC day, before and after the announcement. I focus on the pre-announcement reduction of VIX, which accounts for about 50% of the total decrease around FOMC meetings.

1.1.3 Classification of FOMC announcements via uncertainty reduction

I sort the FOMC days by their reduction of uncertainty during the 24-hour window before announcements into terciles. Figure 1.2 plots the cumulative VIX change and the cumulative return around FOMC meetings for the high-reduction group and low-reduction group, separately. The high-reduction group's VIX index decreases 1.459% significantly over the 2 p.m.-to-2 p.m. pre-FOMC window, which is associated with a deeper pre-FOMC drift (94.4 basis points) than the average FOMC results, as shown in Table 1.2. By contrast, the low-reduction group's VIX index increases instead of decreases before announcements and there is no positive pre-announcement drift.

This classification demonstrates that not all the FOMC announcements are the same—only the ones with uncertainty reduction preceding announcements are associated with the positive pre-FOMC drift. Later I show this is consistent with the full model in section 1.4 that the pre-announcement drift only occurs when the insider is informed.

In addition to the uncertainty reduction prior to FOMC announcements, Abdi and Wu (2018) find that corporate bond returns and trade directions before FOMC announcements predict the pre-FOMC stock market returns. Park (2019) shows that speculators' spread trades in bond futures have predictive information about future FOMC meetings and concludes that private information plays a key role in explaining the pre-FOMC drift. All of these asset-market-evidence indicates the presence of private information before FOMC news.

1.1.4 Measurement of informed trading

Informed trading is not directly observable. Following the microstructure literature, I measure informed trading activity by the order imbalance in the testing security defined as $\frac{B-S}{B+S}$, where B (S) is the aggregate buyer-initiated (seller-initiated) trading volume.[15]

I use two measures of imbalance, OIN and OID, where volume is defined as number of trades and dollar trading volume, respectively.

[15] Ahern (2020) finds order imbalance is one of the most robust predictors of insider trading after all controls.

Bernile, Hu, and Tang (2016) argue that the E-mini S&P 500 futures (E-mini) is the best testing security for the pre-FOMC drift.[16] Following their paper, I classify trading volume of E-mini as buyer- or seller-initiated using the tick rule. Specifically, a transaction is classified as buyer-initiated (seller-initiated) if the transaction price is higher (lower) than the last different transaction price. For each time window, the corresponding order imbalance is the difference between the total buyer- and seller-initiated volumes divided by the total trading volume.

To study the pre-FOMC drift, I examine the 24-hour window preceding FOMC announcements, $[-24H, 0]$. Informed trading leads to the diffusion of private information and uncertainty reduction before announcements. Therefore, for each announcement, I construct a categorical variable, UR, that equals positive one (negative one) when the uncertainty reduces, and the cumulative return is positive (negative) over the 24-hour window. UR is zero otherwise.

In Figure 1.3, for each FOMC announcement, I plot the order imbalance based on number of trades (OIN) and dollar volume (OID) in the 24-hour window before FOMC. When there is uncertainty reduction ($UR = \pm 1$), most order imbalances tend to be in the direction of the realized return before announcements and large in magnitude. While when the uncertainty does not decrease before announcements ($UR = 0$), the order imbalance is smaller and largely random. Table 1.3 compares the average order imbalances in $[-24H, 0]$ window when uncertainty reduces before FOMC announcements ($UR = \pm 1$) and when uncertainty does not reduce before FOMC announcements ($UR = 0$). The average order imbalances are significantly positive on days with pre-FOMC uncertainty reduction, which is consistent with the positive pre-FOMC drift. The difference between the average OIN (OID) of group $UR = \pm 1$ and group $UR = 0$ is 1.99% (2.85%) with a t-stat of 5.11 (5.24) in the 24-hour window before FOMC, which is not only statistically but also economically significant. The trading activity across uncertainty-reduced and non-uncertainty-reduced announcements show notable differences, supporting the presence of informed trading before announcements when there is uncertainty reduction.[17]

[16] Here are their three reasons. First, the asset underlying ES contracts is the S&P 500 index. Second, E-mini is available for trading almost 24 hours on the Globex electronic platform of the CME. Third, E-mini is substantially more liquid comparing to other products.

[17] The same pattern holds for other pre-event windows, such as $[-12H, 0]$ and $[-24H, -12H]$.

Next, I assess the statistical significance of these differences. To measure *abnormal* trading activities on announcement days, I also calculate the order imbalances in the same trading hour windows of non-announcement days in the 21 trading days prior to the current FOMC announcement. I regress the two order imbalance measures, OIN and OID on the announcement indicator (ANN) and the uncertainty-reduced indicator (UR). Table 1.4 reports the ordinary least squares (OLS) coefficient estimates.

The UR coefficient estimates in Columns 1 (OIN) and 2 (OID) are positive, and statistically significant, with t-stat of 5.61 and 4.60, respectively. When there is uncertainty reduction, on average, in the 24-hour window, the number and dollar volume of market orders executed in the direction of the realized pre-FOMC return exceed those in the wrong direction by 1.85% and 2.17% of the total volume, respectively. As shown in Columns 3-6, the similar pattern holds for other pre-event windows, such as $[-24H, -12H]$ and $[-12H, 0]$. It provides robust evidence of informed trading in the 24-hour window before FOMC announcements when there is uncertainty reduction, agreeing with the private information explanation for the pre-FOMC drift.[18]

1.2 The standard asset pricing framework

To provide macroeconomic conditions for the market makers' pricing decisions, in this section, I present a standard consumption-based asset pricing framework that the economy's growth rate is not observable. Later on in section 1.3, I introduce the key elements in the microstructure literature, including the insider, liquidity traders, and the market makers.

[18] Bernile, Hu, and Tang (2016) find evidence consistent with informed trading only until about 30 minutes before scheduled FOMC announcements. After I replicate their paper, I extend their results to $[-24H, 0]$ and find the coefficient estimates of their surprise indicator SUR is not significantly different from zero. The main difference is that we have different definitions of when the insider trading may happen before FOMC announcements. They think the insider trading may only happen when the surprise of FOMC news is large so that the insider can make huge profits. While my explanation is based on the risk-reduction channel instead of unexpected news, the insider can make huge profits from the uncertainty of the FOMC news, even the mean of the news is the same as expected. This is consistent with the model's prediction, as shown in section 1.3.

1.2.1 Physical setup of the model

There are a large number of identical infinitely lived households in the economy. I assume that the consumption of the representative agent, C_t, follows

$$\frac{dC_t}{C_t} = m_t dt + \sigma_C dB_{C,t}, \tag{1.1}$$

where m_t is a continuous-time AR(1) process (an Ornstein-Uhlenbeck process) unobservable to the agent in the economy. The law of motion of m_t is

$$dm_t = a_m \left(\bar{m} - m_t\right) dt + \sigma_m dB_{m,t}. \tag{1.2}$$

The standard Brownian motions $B_{C,t}$ and $B_{m,t}$ in equations (1.1) and (1.2), respectively, are independent.

At time 0, the agent's prior belief about m_0 can be represented by a normal distribution. Although m_t is not directly obeservable, the agent can use two sources of information to update belief about m_t. First, the realized consumption path contains information about m_t, and second, at pre-scheduled discrete time points $T, 2T, 3T, \cdots$, additional signals about m_t are revealed through announcements. For $n = 1, 2, 3, \cdots$, I denote s_n as the signal observed at time nT and assume $s_n = m_{nT} + \varepsilon_n$, where ε_n is i.i.d. over time, and normally distributed with mean zero and variance σ_s^2.

Given the information structure, the posterior distribution of m_t is Gaussian and can be summarized by its first two moments. I define $\hat{m}_t = E_t\left[m_t\right]$ as the posterior mean and $q_t = E_t\left[\left(m_t - \hat{m}_t\right)^2\right]$ as the posterior variance, respectively, of m_t given information up to time t. For $n = 1, 2, \cdots$, at time $t = nT$, the agent updates her belief using Bayes' rule:

$$\hat{m}_{nT}^{+} = q_{nT}^{+}\left[\frac{1}{\sigma_s^2} s_n + \frac{1}{q_{nT}^{-}} \hat{m}_{nT}^{-}\right]; \qquad \frac{1}{q_{nT}^{+}} = \frac{1}{\sigma_s^2} + \frac{1}{q_{nT}^{-}}, \tag{1.3}$$

where $\hat{m}_{nT}^{+}$ and q_{nT}^{+} are the posterior mean and variance after announcements, and $\hat{m}_{nT}^{-}$ and q_{nT}^{-} are the posterior mean and variance before announcements, respectively. A special case is that the announcements can completely reveal the information about m_t, which means, $\sigma_s^2 = 0$. Therefore, σ_s^2 measures the transparency of FOMC announcements.

In the interior of $(nT, (n+1)T)$, the agent updates her belief based on the observed consumption process using the Kalman-Bucy filter:

$$d\hat{m}_t = a_m [\bar{m} - \hat{m}_t] dt + \frac{q(t)}{\sigma_C} d\tilde{B}_{C,t}, \tag{1.4}$$

where the innovation process, $\tilde{B}_{C,t}$ is defined by $d\tilde{B}_{C,t} = \frac{1}{\sigma_C}\left[\frac{dC_t}{C_t} - \hat{m}_t dt\right]$. The posterior variance, $q(t)$ satisfies the Riccati equation:

$$dq(t) = \left[\sigma_m^2 - 2a_m q(t) - \frac{1}{\sigma_C^2} q^2(t)\right] dt. \tag{1.5}$$

1.2.2 Preferences and the SDF

I assume that the representative agent is endowed with a Kreps-Porteus preference with risk aversion γ and intertemporal elasticity of substitution ψ.[19] In continuous time, the preference is represented by a stochastic differential utility, which can be specified by a pair of aggregators $(f, \mathcal{A})$ such that in the interior of $(nT, (n+1)T)$,

$$dV_t = [-f(C_t, V_t) - \frac{1}{2}\mathcal{A}(V_t)||\sigma_V(t)||^2]dt + \sigma_V(t)dB_t \tag{1.6}$$

I adopt the convenient normalization $\mathcal{A}(v) = 0$ and denote $\bar{f}$ the normalized aggregator. Under this normalization, $\bar{f}(C, V)$ is:

$$\bar{f}(C,V) = \frac{\rho}{1 - 1/\psi} \frac{C^{1-1/\psi} - ((1-\gamma)V)^{\frac{1-1/\psi}{1-\gamma}}}{((1-\gamma)V)^{\frac{1-1/\psi}{1-\gamma}-1}}. \tag{1.7}$$

The case of $\psi = 1$ is obtained as the limit of (1.7) with $\psi \to 1$:

$$\bar{f}(C,V) = \rho V [(1-\gamma)\log C - \log[(1-\gamma)V]].$$

Because announcements typically result in discrete jumps in the posterior belief about m_t, the value function is typically not continuous at announcements. Given our normalization of the utility function, for $t = nT$, the pre-announcement utility and post-announcement utility are related by:

$$V_t^- = \mathbb{E}_t^-\left[V_t^+\right],$$

[19] In this paper, I focus on the recursive utility. I can extend it to other preferences that satisfy generalized risk sensitivity defined in Ai and Bansal (2018).

where $\mathbb{E}_t^-$ represents expectation with respect to the pre-announcement information at time t.

In the above setup, I can show that the value function of the representative agent takes the form

$$V(\hat{m}, t, C_t) = \frac{1}{1-\gamma} H(\hat{m}, t) C_t^{1-\gamma},$$

for some twice continuously differentiable function $H(\hat{m}, t)$. The HJB equation and the corresponding boundary conditions for $H(\hat{m}, t)$ can be found in Appendix. Given the utility of the representative agent, the state price density, denoted $\{\pi_t\}_{t=0}^{\infty}$ can be characterized by the following lemma.

Lemma 1. *For $n = 1, 2, 3\cdots$, in the interior of $((n-1)T, nT)$, π_t is a continuous diffusion process with the law of motion*

$$\frac{d\pi_t}{\pi_t} = -r(\hat{m}, t)\, dt - \sigma_\pi(\hat{m}, t)\, d\tilde{B}_{C,t},$$

where $r(\hat{m}, t)$ is the instantaneous risk-free interest rate and $\sigma_\pi(\hat{m}, t)$ is the market price of risk. At announcements, $t = nT$, π_t is discountinuous, and the announcement stochastic discount factor (A-SDF) is given by

$$\Lambda^*_{t,t+\triangle} = \frac{\left[H(\hat{m}_{t+\Delta}, t+\Delta)\right]^{\frac{\frac{1}{\psi}-\gamma}{1-\gamma}}}{\left[\mathbb{E}_t\left(H(\hat{m}_{t+\Delta}, t+\Delta)\right)\right]^{\frac{\frac{1}{\psi}-\gamma}{1-\gamma}}}. \tag{1.8}$$

For convenience, I focus on unit IES $\psi = 1$, which results in

$$H(\hat{m}_t, t) = e^{-\frac{\gamma-1}{a_m+\rho}\hat{m}_t + \mathcal{H}(t)} \equiv e^{-\gamma^A \hat{m}_t + \mathcal{H}(t)},$$

where $\gamma^A \equiv \frac{\gamma-1}{a_m+\rho}$.[20] *Moreover, the A-SDF $\Lambda^*_{t,t+\triangle}$ is counter-cyclical if and only if the agent has early resolution of uncertainty, i.e., $\gamma > \frac{1}{\psi}$, which is equivalent to $\gamma^A > 0$ when $\psi = 1$.*

Given the above A-SDF, any asset with value $A\left(\hat{m}_t^+, t^+\right)$ upon announcements $t = nT$ will be valued by weighting the payoffs through investors' future marginal utility and taking expectations:

[20] *The proof in the appendix provides the formula of SDF for a general IES. All the main results hold under the general IES, which are available upon request.*

$$P\left(\hat{m}_t^-, t^-\right) = \mathbb{E}\left[\frac{H\left(\hat{m}_t^+, t^+\right)}{\mathbb{E}_t^-\left[H\left(\hat{m}_t^+, t^+\right)\right]} A\left(\hat{m}_t^+, t^+\right) | \hat{m}_t^-, q_t^-\right], \qquad t = nT.$$

When there is no private information prior to announcements, the FOMC information is only revealed upon announcements, which results in reductions of uncertainty, and realizations of the equity premium *at*, *rather than ahead of*, the announcements, as shown in Ai and Bansal (2018), Ai, Bansal, Im and Ying (2018), and Wachter and Zhu (2018).

1.3 The benchmark: risk-compensated market makers

To capture the pre-FOMC announcement drift as well as the uncertainty reduction before announcements, I introduce the insider trading into this macroeconomic framework. I extend Kyle's (1985) model (in the continuous-time formulation given by Back (1992)) to allow that market makers are compensated for the risk of assets' fundamentals, where the market makers estimate the discounted value of the risky asset and the A-SDF simultaneously before announcements.

1.3.1 Model setting

The insider in the stock market observes the signal of announcements $s_n = x_{nT} + \varepsilon_n$ at $t = nT - 1$, which happens before FOMC announcements.[21] Thus, she knows the underlying expected growth rate $\hat{m}_{nT}$ and the value of the $A\left(\hat{m}_{nT}, nT\right)$ earlier than other investors in the market. In addition to the insider, there are liquidity traders who have random, price-inelastic demands. All orders are market orders and are observed by all market makers. Denote by Z_t the cumulative orders of liquidity traders through time t. The process Z is assumed to be a Brownian motion independent of ε_n, which has mean zero and variance σ_z^2 (per unit of time). Let X_t denote the cumulative orders of the insider and set $Y = X + Z$.

Given the macroeconomic conditions defined in last section, from equation (1.8), the

[21] In section 1.5.2, I discuss that even the insider is probably informed way before, it is optimal that she starts to trade around the highest average market uncertainty, i.e., 24 hours before announcements shown in Figure 1.1.

market makers' A-SDF at $t = nT - 1$ is:[22]

$$\Lambda^*_{nT-1,nT} = \frac{H\left(\hat{m}_{nT}, nT\right)}{\mathbb{E}_{nT-1}\left[H\left(\hat{m}_{nT}, nT\right)\right]}, \tag{1.9}$$

and later on they update the estimate of the A-SDF based on the observed cumulative order flow before announcements. The market makers, who are competitive, set the price at time $t \in [nT-1,\, nT]$ as

$$P_t = \mathbb{E}\left[\frac{H\left(\hat{m}_{nT}, nT\right)}{\mathbb{E}\left[H\left(\hat{m}_{nT}, nT\right)|\mathcal{F}^Y_t\right]} A\left(\hat{m}_{nT}, nT\right)|\mathcal{F}^Y_t\right], \tag{1.10}$$

where I denote by $\mathcal{F}^Y_t$ the information filtration generated by observing the entire past history of aggregate order flow Y (which I denote by $Y^t = \{Y_s\}_{s\leq t}$). At $t = nT-1$, the market markers have a prior that the expected growth rate upon announcements $\hat{m}_{nT}$ is normally distributed $N\left(\hat{m}_{nT-1}, \Delta Q\right)$ where $\Delta Q = q_{nT-1} - q_{nT}$, as other agents in the economy (except the insider).[23] Here $\frac{1}{q_{nT}} = \frac{1}{\sigma_s^2} + \frac{1}{q_{nT-1}}$ from Bayes' rule. I follow the literature to assume the value of $A\left(\hat{m}_{nT}, nT\right)$ follows a log-normal distribution.[24] More specifically, I specify $\log A\left(\hat{m}_{nT}, nT\right) = \beta \hat{m}_{nT} + N\left(nT\right)$, where $\beta > 0$ measures how the asset value moves with respect to the fundamental.[25]

Given the insider knows the expected growth rate $\hat{m}_{nT}$ at $t = nT-1$, there is no uncertainty of the underlying fundamental to her since then. Thus, $\text{A-SDF}^{\text{insider}}_{t,nT} \equiv 1$ under the insider's information set for all $t \in [nT-1,\, nT]$. In other words, the insider is "risk-neutral" toward the news contained in announcements due to her perfect knowledge of the underlying information.[26] The insider maximizes the expectation of

[22] Similar to Ai and Bansal (2018), I assume that aggregate consumption does not instantaneously respond to the FOMC announcements. This assumption is well motivated because the announcement returns are realized in a 24-hour window before FOMC and the consumption response, if any, at this frequency is not likely to be significant enough to rationalize the magnitude of the premium.

[23] See the proof in Lemma 5.

[24] It can be extended to a general smooth distribution as shown in the proof of section 1.4 in the appendix.

[25] One example of $A\left(\hat{m}_t, t\right) \approx e^{\frac{\phi-1}{a_m + e^{-\bar{\rho}}}\hat{m}_t + N(t)}$ is through the stock which has the claim to the following dividend process:

$$\frac{dD_t}{D_t} = \left[\bar{m} + \phi\left(m_t - \bar{m}\right)\right]dt + \phi\sigma_C dB_{C,t}, \tag{1.11}$$

where we allow the leverage parameter $\phi \geq 1$ so that dividends are more risky than consumption, as in Bansal and Yaron (2004). The proof is in the appendix.

[26] This can be shown directly through equation (1.9) where I take the expectation under the insider's information set at $t \in [nT-1,\, nT]$.

her terminal profit:

$$
\begin{aligned}
&J\left(nT-1, P_{nT-1}, A\left(\hat{m}_{nT}, nT\right)\right) \\
&= \max_{X_t} \mathbb{E}\left[\int_{nT-1}^{nT} \left(A\left(\hat{m}_{nT}, nT\right) - P_t\right) dX_t | \mathcal{F}_{nT-1}^{Y}, A\left(\hat{m}_{nT}, nT\right)\right] \\
&= \max_{\theta_t \in \mathcal{A}} \mathbb{E}\left[\int_{nT-1}^{nT} \left(A\left(\hat{m}_{nT}, nT\right) - P_t\right) \theta_t dt | \mathcal{F}_{nT-1}^{Y}, A\left(\hat{m}_{nT}, nT\right)\right].
\end{aligned} \tag{1.12}
$$

In addition to the entire past history of aggregate order flow Y, the insider knows the actual value of the stock $A\left(\hat{m}_{nT}, nT\right)$, and, of course, her own trading. Following Back (1992), I assume that the insider chooses an absolutely continuous trading rule $dX_t = \theta_t dt$ that belongs to an admissible set $\mathcal{A} = \left\{\theta \text{ s.t. } \mathbb{E}\left[\int_{nT-1}^{nT} \theta_s^2 ds\right] < \infty\right\}$. Therefore, the dynamics of aggregate order flow Y is the sum of the insider's demand the liquidity traders' demand:

$$
dY_t = \theta_t dt + dZ_t.
$$

1.3.2 The equilibrium

Definition 1. *An equilibrium is a price process and an admissible trading strategy, (P_t, θ_t), that satisfy the market makers' rationality condition (1.10) while solving the insider's optimality condition (1.12).*

The introduction of the risk-compensated market makers leads to the following main difference comparing to the standard Kyle model and related extensions in the literature.[27] Instead of only estimating the (discounted) value of the risky asset, the market makers also update the A-SDF simultaneously before announcements based on the observed cumulative order flow. In the standard Kyle model, the risk-neutral market makers only estimate the value of $A\left(\hat{m}_{nT}, nT\right)$ and set the price

$$
P_t^{\text{Kyle}} = \mathbb{E}\left[A\left(\hat{m}_{nT}, nT\right) | \mathcal{F}_t^Y\right], \tag{1.13}
$$

In equilibrium, P_t^{Kyle} must be a martingale under market makers' information set. While when market makers are compensated for the risk of assets' fundamentals, the pricing rule from (1.10) can be rewritten as

$$
P_t = \frac{\mathbb{E}\left[H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right) | \mathcal{F}_t^Y\right]}{\mathbb{E}\left[H\left(\hat{m}_{nT}, nT\right) | \mathcal{F}_t^Y\right]} \equiv \frac{V_t}{\Lambda_t}, \tag{1.14}
$$

[27] See Kyle (1985), Back (1992), Collin-Dufresne and Fos (2016), Back, Crotty, and Li (2018), etc.

where V_t and $\mathit{\Lambda}_t$ are the market makers' estimation of $H(\hat{m}_{nT}, nT)\, A(\hat{m}_{nT}, nT)$ and $H(\hat{m}_{nT}, nT)$, respectively. In equilibrium, both V_t and $\mathit{\Lambda}_t$ are martingales under market makers' information set. However, the pricing rule probably not. Later I will show P_t is a *submartingale* with respect to the market makers' information if and only they are compensated for the risk of assets' fundamentals. The intuition is that the uncertainty of underlying fundamental is resolved gradually after observing aggregate order flow, which is associated with the realization of the premium when the market makers are compensated for risk-taking.

To solve for an equilibrium, I proceed in a few steps. First, in Lemma 2, conditional on a conjectured insider's trading strategy, I derive the stock price dynamics consistent with the market makers' filtering. Then, given the assumed dynamics of the equilibrium price, I solve the insider's optimal trading strategy that is captured in Lemma 3 and Lemma 4. Finally, I show that the conjectured rule by the market makers is indeed consistent with the insider's optimal choice, as stated in Theorem 1.

Lemma 2. $\forall t \in [nT-1,\, nT]$, *suppose the insider adopts the following trading strategy*

$$\theta_t = \frac{\log\left[A(\hat{m}_{nT}, nT)\right] - \mu_P + \gamma^A \beta \Delta Q}{(nT-t)\,\lambda} - \frac{Y_t}{nT-t}, \tag{1.15}$$

where $\mu_P = \beta \hat{m}_{nT-1} + N(nT)$, $\sigma_v^2 = \beta^2 \Delta Q$, *and* $\lambda = \frac{\sigma_v}{\sigma_z}$. *Then the market makers' estimations given by equation (1.14) satisfy the stochastic differential equations*

$$\frac{dV_t}{V_t} = \frac{\beta - \gamma^A}{\beta}\lambda\left[dY_t - \hat{\theta}_t dt\right] \equiv \frac{\beta-\gamma^A}{\beta}\lambda d\hat{Y}_t, \tag{1.16}$$

$$\frac{d\Lambda_t}{\Lambda_t} = \frac{-\gamma^A}{\beta}\lambda\left[dY_t - \hat{\theta}_t dt\right] \equiv \frac{-\gamma^A}{\beta}\lambda d\hat{Y}_t. \tag{1.17}$$

The expected insider's order rate under the market makers' filtration $\mathcal{F}_t^Y$ *is*

$$\hat{\theta}_t \equiv \mathbb{E}\left[\theta_t | \mathcal{F}_t^Y\right] = \frac{\gamma^A \beta \Delta Q}{\lambda}, \tag{1.18}$$

and the adjusted order flow $\hat{Y}_t$

$$\hat{Y}_t \equiv Y_t - \int_{nT-1}^{t} \hat{\theta}_s\, ds = Y_t - \frac{\gamma^A \beta \Delta Q}{\lambda}\left[t - (nT-1)\right], \tag{1.19}$$

is a Brownian Motion with instant variance σ_z^2 *with respect to the market makers' filtration* $\mathcal{F}_t^Y$.

Further, the market makers' pricing rule in equation (1.14) is a function of $\left(t, \hat{Y}_t\right)$ *that follows*

$$\frac{dP\left(t, \hat{Y}_t\right)}{P\left(t, \hat{Y}_t\right)} = \lambda d\hat{Y}_t + \gamma^A \beta \Delta Q dt, \quad \textit{with } P_{nT-1} = e^{\mu_P - \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2 + \frac{1}{2}\left(\frac{\beta - \gamma^A}{\beta}\right)^2 \sigma_v^2}. \quad (1.20)$$

When market makers are risk neutral, aggregate order flow Y_t at equilibrium is a martingale under the market makers' information set, as shown by Back (1992). In other words, the market makers are to set the pricing rule such that the expected order rate from the insider is zero. While market makers are risk averse to the underlying fundamental (i.e., $\gamma^A > 0$), equation (1.18) indicates that the expected insider's order rate under the market makers' filtration $\mathcal{F}_t^Y$ is strictly positive. Here is the intuition behind this result. The information from aggregate order flow resolves market makers' uncertainty before FOMC announcements. Since they are compensated for risk-taking, the equity premium realized gradually during this period. This leads to an average upward drift in market prices. Therefore, market makers would expect an average positive trading volume from the insider to chase that premium. Besides, the insider has to consider this additional price impact from uncertainty resolution when they trade, which is unique in this model.

The above analysis implies aggregate order flow Y_t at equilibrium is no longer a martingale under $\mathcal{F}_t^Y$ when $\gamma^A > 0$. More importantly, since the average positive order flow from the insider is expected, market markers would update their estimates from the adjusted order flow $\hat{Y}_t$ instead of aggregate order flow Y_t. Thus, I suppose that there exists an equilibrium with two state variables: time t and the adjusted order flow $\hat{Y}_t$. Then given the make makers' pricing rule, $P(t) = P\left(t, \hat{Y}_t\right)$, the insider chooses the order rate to maximize her trading profit. That is,

$$J\left(t, y, A\left(\hat{m}_{nT}, nT\right)\right) = \max_{\theta_t \in \mathcal{A}} \mathbb{E}\left[\int_t^{nT} \left(A\left(\hat{m}_{nT}, nT\right) - P\left(s, \hat{Y}_s\right)\right) \theta_s ds | \hat{Y}_t = y, A\left(\hat{m}_{nT}, nT\right)\right]$$

subject to

$$d\hat{Y}_t = \left[\theta_t - \hat{\theta}_t\right] dt + dZ_t, \quad \text{where } \hat{\theta}_t \equiv \mathbb{E}\left[\theta_t | \mathcal{F}_t^Y\right]. \quad (1.21)$$

The principle of optimality implies the following Bellman equation

$$\max_{\theta_t \in \mathcal{A}} \left\{ \left(A\left(\hat{m}_{nT}, nT\right) - P\left(t, y\right)\right) \theta_t + J_t + J_y \left[\theta_t - \hat{\theta}_t\right] + \frac{1}{2}\sigma_z^2 J_{yy} \right\} = 0, \quad (1.22)$$

where the subscripts denote the derivatives. The necessary conditions for having an optimal solution to the Bellman equation (1.22) are

$$\begin{aligned} J_y\left(t, y, A\left(\hat{m}_{nT}, nT\right)\right) &= P\left(t, y\right) - A\left(\hat{m}_{nT}, nT\right), \quad (1.23) \\ J_t + \frac{1}{2}\sigma_z^2 J_{yy} - \hat{\theta}_t J_y &= 0. \quad (1.24) \end{aligned}$$

These necessary conditions lead to the following results.

Lemma 3. *Suppose the expected order rate* $\hat{\theta}\left(t\right) = \Theta\left(t, \hat{Y}_t\right)$, *where* $\hat{Y}_t$ *is the adjusted order at* t. *Let* $\omega_t = y$ *and suppose that the stochastic differential equation*

$$d\omega_s = dZ_s - \Theta\left(s, \omega_s\right) ds, \quad \forall nT \geq s \geq t \geq nT - 1$$

has a unique solution, where Z_s *is a Brownian motion with instant variance* σ_z^2.[28] *If there exists a strictly monotone function* $g\left(\cdot\right)$ *such that the pricing rule is*

$$P\left(t, y\right) = \mathbb{E}\left[g\left(\omega_{nT}\right) \middle| \omega_t = y\right], \quad (1.25)$$

then

$$J\left(t, y, A\left(\hat{m}_{nT}, nT\right)\right) = \mathbb{E}\left[j\left(\omega_{nT}, A\left(\hat{m}_{nT}, nT\right)\right) \middle| \omega_t = y\right], \quad (1.26)$$

is a smooth solution to the Bellman equations (1.23) and (1.24), where

$$j\left(y, A\left(\hat{m}_{nT}, nT\right)\right) = \int_y^{g^{-1}\left(A\left(\hat{m}_{nT}, nT\right)\right)} \left[A\left(\hat{m}_{nT}, nT\right) - g\left(x\right)\right] dx \geq 0, \quad \forall \left(y, A\left(\hat{m}_{nT}, nT\right)\right).$$

Lemma 4. *Any continuous trading strategy that makes* $\lim_{t \to nT} P\left(t, \hat{Y}\left(t\right)\right) = A\left(\hat{m}_{nT}, nT\right)$ *is optimal, where* $P\left(t, y\right)$ *is as defined by equation (1.25).*

Having established these results, I can now proceed to characterize the equilibrium price and the insider's optimal strategy. The equilibrium I obtain, which constitutes the main results of this paper, is summarized in the following theorem.

Theorem 1. $\forall t \in \left[nT - 1, \, nT\right]$, *there exists an equilibrium where the price process* P_t *and optimal strategy of the insider* θ_t *have dynamics,*

$$\begin{aligned} \frac{dP\left(t, \hat{Y}_t\right)}{P\left(t, \hat{Y}_t\right)} &= \lambda d\hat{Y}_t + \gamma^A \beta \Delta Q dt, \quad (1.27) \\ \theta\left(t, \hat{Y}_t\right) &= \frac{\log\left[A\left(\hat{m}_{nT}, nT\right)\right] - \mu_P}{\left(nT - t\right)\lambda} - \frac{\hat{Y}_t}{nT - t} + \frac{\gamma^A \beta \Delta Q}{\lambda}, \quad (1.28) \end{aligned}$$

[28] *Note that* ω_t *is the adjusted order when the insider does not submit orders.*

where $\hat{Y}_t$, P_{nT-1}, μ_P, σ_v, and λ are defined in Lemma 2. The expected insider's order rate under $\mathcal{F}_t^Y$ is defined in equation (1.18).

The maximized expected profit of the insider is

$$J\left(t, P\left(t, \hat{Y}_t\right), A\left(\hat{m}_{nT}, nT\right)\right) = \frac{1}{2}\sigma_z\sigma_v\left(nT - t\right)A\left(\hat{m}_{nT}, nT\right)$$
$$+\frac{P\left(t, \hat{Y}_t\right) - A\left(\hat{m}_{nT}, nT\right) + A\left(\hat{m}_{nT}, nT\right)\left[\log A\left(\hat{m}_{nT}, nT\right) - \log P\left(t, \hat{Y}_t\right)\right]}{\lambda}. \tag{1.29}$$

With respect to the insider's filtration, $P\left(t, \hat{Y}_t\right)$ converges almost surely to $A\left(\hat{m}_{nT}, nT\right)$ at time $t = nT$. When the market makers are risk-compensated, with respect to the market makers' filtration, both the pricing rule $P\left(t, \hat{Y}_t\right)$ and the price-response coefficient $P_{\hat{Y}}\left(t, \hat{Y}\right)$ are submartingales with a constant growth rate $\gamma^A\beta\Delta Q$.

Further, $\forall t \in [nT-1, nT]$, the expected cumulative pre-FOMC announcement drift is

$$\log \mathbb{E}\left[\frac{P_t}{P_{nT-1}}|\mathcal{F}_{nT-1}^Y\right] = \gamma^A\beta\Delta Q\left(t - \left(nT - 1\right)\right). \tag{1.30}$$

This implies there is a strictly positive pre-FOMC announcement drift if and only if the market makers are compensated for the risk of assets' fundamentals, i.e., $\gamma^A > 0$.

I now comment on several implications of the theorem. First, the equilibrium price is a submartingale, expected to increase over time. This contrasts my framework from much of the literature. They find the price dynamics is a martingale under the risk-neutral market makers since they are indifferent, to resolve the uncertainty now or in the future. While when market makers are risk compensated, the resolution of uncertainty is associated with the realizations of the premium. The positive expected pre-FOMC announcement premium is cumulated at a constant rate $\gamma^A\beta\Delta Q$, which is the negative covariance between the innovation to the A-SDF and the asset value. Intuitively, the pre-annoucement drift would be larger: (1) when market mares are more risk-averse to the underlying fundamental; (2) the asset value has a lager exposure to the FOMC news; (3) more transparent FOMC announcements which reduce more uncertainty. In addition, the equilibrium price converges to the value $A\left(\hat{m}_{nT}, nT\right)$, known ex ante only to the insider, at FOMC announcements. This guarantees all of the private information is eventually incorporated into the price and generalizes the result proved in Back (1992) under the risk-neutral market makers.

Second, I find when the market makers are risk averse to the underlying fundamental (i.e., $\gamma^A > 0$), the expected insider's order rate under the market makers' filtration $\mathcal{F}_t^Y$ follows

$$\mathbb{E}\left[\theta_t|\mathcal{F}_t^Y\right] = \frac{\gamma^A \beta \Delta Q}{\lambda} = \gamma^A \sqrt{\Delta Q}\sigma_z, \tag{1.31}$$

which is strictly positive. This is very different form the Kyle model, where the expected insider's order rate is always zero. Here is the intuition. Due to the average upward drift in market prices, the market makers rationally anticipate that the insider would trade positively on average to chase that premium. The insider also has to consider the additional price impact from uncertainty resolution when she trades, which is unique in this model. The equilibrium expected insider's order rate is determined by the ratio of the expected pre-FOMC announcement premium per unit of time ($\gamma^A \beta \Delta Q$) to Kyle's lambda (λ). Therefore, the abnormal order imbalances are on average positive when there is private information before FOMC announcements. In addition, the insider would on average trade more aggressively when market mares are more risk-averse to the uncertainty or FOMC announcements are more transparent, which is caused by the higher realized equity premium per unit of time. In the meantime, when noise traders are more active, the insider on average trades more due to the smaller price impact, which has been largely missed in the Kyle-type models.[29]

Third, the price impact $P_{\hat{Y}}\left(t, \hat{Y}\right)$ is a also submartingale, which grows at the same rate the equilibrium price. The risk-averse market makers benefit from uncertainty resolution out of observing aggregate order flow. Therefore, to entice the insider to trade and release information early, the market makers have incentives to set the price impact that increases on average. Collin-Dufresne and Fos (2016) is one of the few papers that achieve the same result through a different channel that comes from the insider's potential benefit to wait for better liquidity with stochastic noise trading volatility.[30]

[29] The only exception is Collin-Dufresne and Fos (2016) that derive the same result by assuming noise trading volatility follows a general stochastic process.

[30] The price impact is constant in Kyle (1985). In extensions of that model Back (1992), Back and Pedersen (1998), Baruch (2002), Back and Baruch (2004), Caldentey and Stacchetti (2010), price impact is either a martingale or a supermartingale. Collin-Dufresne and Fos (2016) points that Foster and Viswanathan (1996) and Back, Cao, and Willard (2000) may also generate an increase in the deterministic price impact, at least near the end of the trading horizon, because of competition among multiple informed traders.

Fourth, when the market makers converge to be risk-neutral, the limit of the equilibrium is well defined and converges to the traditional Kyle model (more precisely, converges to Back (1992)). When γ^A converges to zero, the expected insider's order rate $\hat{\theta}_t$ converges to zero, which implies aggregate order flow Y_t converges to a martingale. The pricing rule $P\left(t, \hat{Y}_t\right)$ and the price-response coefficient $P_{\hat{Y}}\left(t, \hat{Y}\right)$ also converge to martingales. This implies the expected pre-announcement drift converges to a flat line, as in Back (1992). The convergence result demonstrates that there is a strictly positive pre-FOMC announcement drift if and only if the market makers are compensated for the risk of assets' fundamentals, i.e., $\gamma^A > 0$, as proved in Theorem 1.

1.3.3 Properties of equilibrium

Having characterized the equilibrium, in this section, I study the equilibrium properties and map the model to asset market fluctuations before FOMC announcements.

The following proposition captures the uncertainty reduction prior to announcements in the equilibrium from Theorem 1.

Proposition 1. *With respect to the market makers' filtration $\mathcal{F}_t^Y$, $\forall t \in [nT-1, nT]$, the uncertainty reduction at time t comparing to $nT-1$ follows*

$$Var\left[\log P_{nT}|\mathcal{F}_t^Y\right] - Var\left[\log P_{nT}|\mathcal{F}_{nT-1}^Y\right] = -\beta^2 \Delta Q\left[t-(nT-1)\right]. \tag{1.32}$$

Thus, prior to announcements, the uncertainty reduces at a constant rate $\beta^2\Delta Q$ per unit of time.

So far, $\forall t \in [nT-1, nT]$, I explicitly characterize the expected pre-FOMC announcement drift and the implied variance reduction in equations (1.30) and (1.32), respectably. I take the macroeconomic parameters from Ai and Bansal (2018) and calibrate the risk aversion and the transparency of announcements to match the level of cumulative return and uncertainty reduction upon announcements.[31] The parameters are reported in Table 1.1. I call this case as the benchmark where the market makers are risk-compensated. For comparison, I study another case where I keep other parameters the same and assume the market makers are risk-neutral, which is equivalent to the original Kyle model with a log-normal distribution of the asset value (see Back (1992)).

[31] See the full model calibration in section 1.4.3.

Figure 1.4 depicts the model implications for the benchmark case (the dotted red line) and the Kyle case (the dashed blue line) as a function of time, respectively. Panel A plots the implicated variance changes before announcements defined in equation (1.32), which are the same for both cases.[32] This is because the implicated variance reduction only depends on the risk exposure β, which is not a function of γ^A. However, the expected pre-FOMC announcement excess returns are very different, as shown in Panel B. When the market makers are risk-compensated, the expected pre-announcement drift has a constant positive rate as captured in equation (1.30). The expected pre-announcement drift is literally zero with the risk-neutral market makers. Panel C compares the expected insider's order rate under the market makers' information set, which is strictly positive with the risk-averse market makers as captured in equation (1.31). Panel D plots the average realized pre-FOMC announcement excess returns that are computed from parallel simulations, respectively. This is consistent with my previous discussion that there is a strictly positive pre-FOMC announcement drift if and only if the market makers are compensated for the risk of assets' fundamentals.

I follow the empirical procedure—the tick rule to calculate the order imbalances in the model. The aggregate buyer-initiated (OI_B) and seller-initiated (OI_S) dollar trading volumes are defined as

$$OI_B = \int_{nT-1}^{nT} p_t \theta_t 1_{\{p_t > p_{t-1}\}} \, dt, \qquad OI_S = \int_{nT-1}^{nT} p_t \theta_t 1_{\{p_t < p_{t-1}\}} \, dt, \tag{1.33}$$

respectively. Thus, the order imbalances are $\frac{OI_B - OI_S}{OI_B + OI_S}$. Figure 1.5 shows the distribution of the order imbalances for the benchmark and the Kyle model. The average of the order imbalances in the benchmark is 0.85%, which is statistically significant at the 1% level. The significant positive order imbalances come from the positive excess return before FOMC announcements. While in the Kyle model, the average is not significantly different from zero that is not consistent with data.

So far, I compare the proprieties of average variables in the benchmark to the Kyle model. To have a better understanding of the equilibrium in Theorem 1, the following proposition emphasizes the differences between the two cases under any realized path of Z_t.

[32] For convenience, I match the level of the cumulative change of the VIX^2 instead of the VIX displayed in Figure 1.1.

Proposition 2. *Under the same realization path of Z_t, the difference between the price P_t in the benchmark ($\gamma^A > 0$) and the price P_t^{Kyle} in Kyle ($\gamma^A_{Kyle} = 0$) increases at a constant rate:*

$$\log P_t - \log P_t^{Kyle} = \gamma^A \beta \Delta Q \left[t - nT\right], \quad \forall t \in \left[nT - 1,\, nT\right],$$

which converges to zero upon announcements. Meanwhile, the difference between the aggregate cumulative trading flow Y_t and Y_t^{Kyle} increases at a constant rate:

$$Y_t - Y_t^{Kyle} = \frac{\gamma^A \beta \Delta Q}{\lambda} \left[t - (nT - 1)\right], \quad \forall t \in \left[nT - 1,\, nT\right].$$

Figure 1.6 plots the log price dynamics and aggregate trading volume for the benchmark case (the dotted red line) and the Kyle case (the dashed blue line) under one realized path of Z_t.[33] Panel A and B plot the dynamics of the log price and the aggregate trading flow as a function of time, repetitively. The two cases show similar patterns of fluctuations, which are consistent with the linear differences in Panel C and D. When the market makers are risk-averse to the underlying fundamental, the initial price P_{nT-1} has to be lower to attract them to hold the assets. The difference of the price converges to zero since all of the private information is eventually incorporated into the price for both cases, i.e., $P_{nT} = P_{nT}^{\text{Kyle}} = A\left(\hat{m}_{nT}, nT\right)$ almost surely. The insider in the benchmark trades more aggressively to chase the realized premium, which results in the larger aggregate trading volume over time.

1.4 Full model: the insider may not be informed

Figure 1.2 indicates that not all FOMC announcements are the same—some are not associated with uncertainty reduction prior to announcements. Motivated by this fact, I extend the above model to the case that the insider may or may not be informed of the signal s_n before announcements.[34] In the meantime, the market makers are not sure whether the insider observes the signal or not. The market makers share a common belief that such an event, in which the insider observes this information earlier than the

[33] Note that although Figure 1.6 plots only one realized path of Z_t, it represents the typical situation in the model.

[34] This extension is based on Li (2013), which extends Back (1992) to study the insider trading with uncertain informed trading. He keeps the assumption of the risk-neutral market makers.

public, occurs with a probability $\pi_{nT-1} \in (0,1)$ at time 0. Therefore, in addition to the discounted value of the risky asset and the A-SDF, the market makers also have to update their estimate of the probability that the insider has private information of FOMC announcements.

1.4.1 Model setting

Let $X_{\delta,t}$ denote the net orders from the insider trader. Then the total cumulative order flow Y_t can be expressed as

$$Y_t = X_{\delta,t} + Z_t,$$

where δ is an indicator function, which is equal to 1 if the insider has information and is equal to 0 otherwise. By observing this order flow, the market makers update their estimates about the *probability* that the insider possesses private information and the *value* of the risky security. Let $\mathcal{F}_{0,t} = \mathcal{F}_t^Y \times \{\delta = 0\}$ under the hypothesis $\delta = 0$ and $\mathcal{F}_{1,t} = \mathcal{F}_t^Y \times \{\delta = 1\}$ under the hypothesis $\delta = 1$. I let $\pi(t) = \mathbb{E}\left[\delta|\mathcal{F}_t^Y\right]$ be the estimate of the probability that the insider has private information at time t.

If the insider does not have any private information ($\delta = 0$), she has no information other than what the market makers have. Therefore, $\forall t \in [nT-1,\, nT]$, the best estimate of the security's value is

$$\begin{aligned} \bar{v}^* &\equiv \mathbb{E}\left[\frac{H(\hat{m}_{nT}, nT)}{\mathbb{E}\left[H(\hat{m}_{nT}, nT)|\mathcal{F}_{0,t}\right]} A(\hat{m}_{nT}, nT)|\mathcal{F}_{0,t}\right], \\ &= \frac{\mathbb{E}\left[H(\hat{m}_{nT}, nT) A(\hat{m}_{nT}, nT)|\mathcal{F}_{0,t}\right]}{\mathbb{E}\left[H(\hat{m}_{nT}, nT)|\mathcal{F}_{0,t}\right]} \equiv \frac{\bar{V}}{\bar{\Lambda}}, \end{aligned} \tag{1.34}$$

where I define $\bar{V}$ and $\bar{\Lambda}$ as the estimate of $H(\hat{m}_{nT}, nT) A(\hat{m}_{nT}, nT)$ and $A(\hat{m}_{nT}, nT)$ under the case that the insider is not informed, respectively.

If the insider has private information ($\delta = 1$), the value estimate of the risky security at time t conditional on $\delta = 1$ is

$$\begin{aligned} v^*(t) &\equiv \mathbb{E}\left[\frac{H(\hat{m}_{nT}, nT)}{\mathbb{E}\left[H(\hat{m}_{nT}, nT)|\mathcal{F}_{1,t}\right]} A(\hat{m}_{nT}, nT)|\mathcal{F}_{1,t}\right], \\ &= \frac{\mathbb{E}\left[H(\hat{m}_{nT}, nT) A(\hat{m}_{nT}, nT)|\mathcal{F}_{1,t}\right]}{\mathbb{E}\left[H(\hat{m}_{nT}, nT)|\mathcal{F}_{1,t}\right]} \equiv \frac{V(t)}{\Lambda(t)}, \end{aligned} \tag{1.35}$$

where I define $V(t)$ and $\Lambda(t)$ as the estimate of $H(\hat{m}_{nT}, nT)\, A(\hat{m}_{nT}, nT)$ and $A(\hat{m}_{nT}, nT)$ under the case that the insider is informed, respectively.

With the uncertainty of δ, the market makers estimate the discounted value under the information structure $\mathcal{F}_{1,t}$ and estimate the probability that the insider has observed private information under the information structure $\mathcal{F}_t^Y$. Given these two estimates, the market makers set the price that follows

$$\begin{aligned} P(t) &= \mathbb{E}\left[\frac{H(\hat{m}_{nT}, nT)}{\mathbb{E}\left[H(\hat{m}_{nT}, nT)\,|\mathcal{F}_t^Y\right]} A(\hat{m}_{nT}, nT)\,|\mathcal{F}_t^Y\right], \\ &= \frac{\mathbb{E}\left[H(\hat{m}_{nT}, nT)\, A(\hat{m}_{nT}, nT)\,|\mathcal{F}_t^Y\right]}{\mathbb{E}\left[H(\hat{m}_{nT}, nT)\,|\mathcal{F}_t^Y\right]}, \\ &= \frac{\pi(t)\, V(t) + (1-\pi(t))\,\bar{V}}{\pi(t)\,\Lambda(t) + (1-\pi(t))\,\bar{\Lambda}}. \end{aligned} \tag{1.36}$$

Note that when the market makers know the insider is always informed ($\pi_{nT-1} = 1$), the market makers set the price as

$$P(t) = \mathbb{E}\left[\frac{H(\hat{m}_{nT}, nT)}{\mathbb{E}\left[H(\hat{m}_{nT}, nT)\,|\mathcal{F}_{1,t}\right]} A(\hat{m}_{nT}, nT)\,|\mathcal{F}_{1,t}\right],$$

which goes back to the benchmark model in section 1.3.

I impose the following restriction on the market makers' value estimates $V(t)$ and $\Lambda(t)$ conditional on $\delta = 1$ defined in equation (1.35):

$$\mathbb{E}\left[\int_{nT-1}^{nT} V^2(s)\, ds\right] < \infty; \quad \mathbb{E}\left[\int_{nT-1}^{nT} \Lambda^2(s)\, ds\right] < \infty.$$

This restriction implies that the pricing rule defined by equation (1.36) satisfies

$$\mathbb{E}\left[\int_{nT-1}^{nT} P^2(s)\, ds\right] < \infty,$$

which is sufficient to rule out the so-called doubling strategy that the insider could use.[35]

1.4.2 The equilibrium

Definition 2. *An equilibrium is an quadruple* (X_0, X_1, P, Π) *such that*

[35] See Back (1992) and Li (2003) for more details.

1. *both* X_0 *and* X_1 *are the optimal trading strategies of the insider when she has not or has observed private information, respectively, given* $P(t)$ *and* Π*;*

2. $P(t) = \frac{\Pi(t)V(t)+(1-\Pi(t))\bar{V}}{\Pi(t)\Lambda(t)+(1-\Pi(t))\bar{\Lambda}}$ *is the stock price at time* t*, where* $V(t)$ *and* $\Lambda(t)$ *are the market makers' value estimates of the risky security and SDF conditional on* $\delta = 1$*, and* $\Pi(t) = \pi(t)$*, is the market makers' probability estimates that the insider has private information, given the insider trader's trading strategies* X_0 *and* X_1*.*

As the benchmark in section 1.3, I assume that the insider chooses an absolutely continuous trading rule

$$dX_{1,t} = \theta\left(t, \tilde{V}\right) dt.$$

$\theta\left(t, \tilde{V}\right)$ belongs to an admissible set $\mathcal{A} = \left\{\theta \text{ s.t. } \mathbb{E}\left[\int_{nT-1}^{nT} \theta^2\left(t, \tilde{V}\right) ds\right] < \infty\right\}$, where $\tilde{V}$ is the insider's perfect knowledge of $H(\hat{m}_{nT}, nT)\, A(\hat{m}_{nT}, nT)$.[36] When the insider has no information other than what the market makers have, her order rate becomes $\theta\left(t, \bar{V}\right)$. Given this trading strategy, the cumulative flow is

$$Y_t = \int_{nT-1}^{t} \theta(s, V)\, ds + Z_t. \tag{1.37}$$

From the market makers' point of view, the cumulative order flow has two possible interpretations because they don't know how much the liquidity traders trade. One is

$$dY_t = \theta\left(t, \tilde{V}\right) dt + dZ_t, \tag{1.38}$$

if the insider is informed, and the other is

$$dY_t = \theta\left(t, \bar{V}\right) dt + dZ_t, \tag{1.39}$$

if the insider is not informed.

The following assumption imposes that when the insider is not informed, she will not take a dramatically different trading strategy. Otherwise, her trading behavior may

[36] Note that the order rate of the insider should also depend on the market makers' pricing rule or some other state variable(s). I omit such state variables in the expression of the order rate because what these variables are is not clear yet. Besides, both of the estimation of $H(\hat{m}_{nT}, nT)\, A(\hat{m}_{nT}, nT)$ and $A(\hat{m}_{nT}, nT)$ rely on the belief updates of $\hat{m}_{nT}$. In other words, inferring one of $V(t)$ and $\Lambda(t)$ is enough. Therefore, I write $\theta\left(t, \tilde{V}\right)$ as a function of t and $V(t)$.

immediately reveal that she does not have private information for a specific FOMC announcement.[37]

Assumption 1. *When the insider is not bettered informed, she maximizes the following terminal profit under her best estimation of the asset value,*

$$\int_{nT-1}^{nT}\left(\mathbb{E}\left[\frac{H\left(\hat{m}_{nT},nT\right)}{\mathbb{E}\left[H\left(\hat{m}_{nT},nT|\mathcal{F}_{nT-1}\right)\right]}A\left(\hat{m}_{nT},nT\right)|\mathcal{F}_{nT-1}\right]-P_s\right)\theta_s ds$$
$$=\int_{nT-1}^{nT}\left(\bar{v}^*-P_s\right)\theta_s ds.$$

Given the observation of the cumulative order flow, the market makers update the probability that the insider has private information, the A-SDF, as well as the discounted value of the security conditional on the insider is informed. These estimates are done by solving a nonlinear filtering problem. The equilibrium is summarized in the following theorem (proved in the Appendix).

Theorem 2. *Under Assumption 1,* $\forall t\in[nT-1,nT]$*, there exists an equilibrium* (X_0,X_1,P,Π) *that follows:*

(1). The market makers' probability estimate $\Pi(t,y)$ *is*

$$\Pi\left(t,y\right)=\frac{\pi_{nT-1}\exp\left(\frac{1}{2\sigma_z^2}\frac{[y-\bar{y}]^2}{nT-t}+\frac{1}{2}\log\left(nT-t\right)-\frac{\bar{y}^2}{2\sigma_z^2}\right)}{1-\pi_{nT-1}+\pi_{nT-1}\exp\left(\frac{1}{2\sigma_z^2}\frac{(y-\bar{y})^2}{nT-t}+\frac{1}{2}\log\left(nT-t\right)-\frac{\bar{y}^2}{2\sigma_z^2}\right)},\tag{1.40}$$

where y *represents the adjusted order flow* $\hat{Y}_{1,t}$ *(defined later) and* $\bar{y}=\frac{\frac{\beta-\gamma^A}{\beta}\sigma_v^2}{2\lambda}$*;*

(2). The pricing rule $P(t,y)$ *has dynamics*

$$P\left(t,y\right)=P_{nT-1}\frac{\Pi\left(t,y\right)e^{\frac{\beta-\gamma^A}{\beta}\lambda y-\frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2\sigma_v^2(t-(nT-1))}+1-\Pi\left(t,y\right)}{\Pi\left(t,y\right)e^{-\frac{\gamma^A}{\beta}\lambda y-\frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2\sigma_v^2(t-(nT-1))}+1-\Pi\left(t,y\right)},\tag{1.41}$$

where P_{nT-1}*,* σ_v*, and* λ *are defined in Lemma 2;*

(3). The insider's trading strategy $X_\delta(t,y)$ *satisfies*

$$X_1\left(t,y\right)=\int_{nT-1}^{t}\theta\left(s,y;H\left(\hat{m}_{nT},nT\right)A\left(\hat{m}_{nT},nT\right)\right)ds,\quad X_0\left(t,y\right)=\int_{nT-1}^{t}\theta\left(s,y;\bar{V}\right)ds.\tag{1.42}$$

[37] It is possible that the insider observes a private signal, indicating that the terminal value is $\bar{v}^*$. Under this case, since the insider has no information advantage comparing to market makers' prior, I interpret it as the insider having no private information.

The insider's order rate for any $\tilde{V}$ is

$$\theta\left(t, y; \tilde{V}\right) = \bar{\theta}\left(t, y\right) + \frac{\left(\log \tilde{V} - \mu_V\right) / \left(\frac{\beta - \gamma^A}{\beta}\lambda\right) - \bar{y} - \Pi\left(t, y\right)\left[y - \bar{y}\right]}{nT - t},$$

where $\mu_V = \left(\beta - \gamma^A\right)\hat{m}_{nT-1} + \mathcal{H}\left(nT\right) + N\left(nT\right)$ *is the mean of* $\log\left[H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right)\right]$.

The expected order rate of the insider $\bar{\theta}\left(t, y\right)$ *under the market makers' filtration* $\mathcal{F}_t^Y$ *satisfies*

$$\bar{\theta}\left(t, y\right) \equiv \mathbb{E}\left[\theta\left(t, y; \tilde{V}\right) | \mathcal{F}_t^Y\right] = \frac{\frac{\gamma^A \beta \Delta Q}{\lambda}\Pi\left(t, y\right) E\left(t, y\right) - \Pi\left(t, y\right)\left(1 - \Pi\left(t, y\right)\right)\frac{y - \bar{y}}{nT - t}\left(E\left(t, y\right) - 1\right)}{\Pi\left(t, y\right) \cdot E\left(t, y\right) + 1 - \Pi\left(t, y\right)}, \tag{1.43}$$

where $E\left(t, y\right) = e^{-\frac{\gamma^A}{\beta}\lambda y - \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2 (t - (nT - 1))}$.

The adjusted order flow $\hat{Y}_{1,t}$ *starts from 0 and follows*

$$d\hat{Y}_{1,t} = \frac{\left(\log \tilde{V} - \mu_V\right) / \left(\frac{\beta - \gamma^A}{\beta}\lambda\right) - \hat{Y}_{1,t}}{nT - t} dt + dZ_t, \tag{1.44}$$

where $\tilde{V} = H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right)$ *when informed and* $\tilde{V} = \bar{V}$ *when not informed.*

The above theorem shows that the benchmark's main results still hold when I extend the model with the potential better-informed insider.[38] Conditional on the insider is better informed, the equilibrium price dynamics is a submartingale when the maker makers are risk-compensated. Both the growth rate of the expected pre-FOMC announcement drift and the expected insider's order rate are time-varying, which are caused by the dynamics of the probability estimate. In addition, the pricing rule is nonlinear and stochastic, which drives price volatility, market depth, and price response to be stochastic.

1.4.3 Properties of equilibrium and model calibration

The stochastic pricing rule in Theorem 2 gives me the hope to match the nonlinear pre-FOMC announcement drift in the data. In this section, I study the equilibrium properties and calibrate the model to the pre-FOMC drift that occurs 24 hours before announcements.

[38] When the insider is always better informed, i.e., $\pi_{nT-1} = 1$, the equilibrium goes back to the benchmark model in section 1.3. When the market makers are risk-neutral, the equilibrium goes back to Li (2013).

Proposition 3. *For any smooth distribution of the prior $G(\pi_{nT-1})$, the average realized pre-FOMC announcement drift just before announcements $(t = nT^-)$ is*

$$\log \mathbb{E}\left[\frac{P_{nT-}}{P_{nT-1}}\right] = \eta\gamma^A\beta\Delta Q,$$

where the expectation is taken over all states of natural and η is the fraction of insider that is informed. Meanwhile, the average uncertainty reduction just before announcements is

$$\mathbb{E}\left[Var\left[\log P_{nT-}|\mathcal{F}^Y_{nT-}\right] - Var\left[\log P_{nT-}|\mathcal{F}^Y_{nT-1}\right]\right] = -\eta\beta^2\Delta Q.$$

Here the expectation is also taken over all states of natural.

Proposition 3 captures the average realized pre-FOMC drift and the average uncertainty reduction just before announcements in the presence of the potential better-informed insider. The intuition is as follows. When the insider is not informed, the market makers figure that out just before announcement (i.e., $\lim_{t\to nT^-}\Pi\left(t, \hat{Y}_{1,t}\right) = 0$) and the price P_{nT^-} equals P_{nT-1} almost surely. Besides, there is no uncertainty reduction since the insider has no information other than what the market makers have at $t = nT - 1$. While when the insider is better informed, all of the private information is eventually incorporated into the price, which is associated with uncertainty reduction. The probability estimate converges to 1 and the price converges to $A(\hat{m}_{nT}, nT)$ almost surely upon announcements.

The closed-form solutions in Theorem 2 and Proposition 3 generate a precise mapping from the model's parameters to the asset market evidence before FOMC announcements. The calibration is summarized in Table 1.1. I choose the fraction of insider that is informed η across these FOMC announcements to be 0.5, consistent with Figure 1.2. Besides, I assume the market makers prior $\pi_0 \equiv 0.2$ to match the nonlinear trend of the pre-FOMC announcement drift.[39] I set $\sigma_m = 0.43\%$ to match the level of the average cumulative pre-FOMC announcement excess return. All other parameters are the same as the benchmark.

Figure 1.7 plots the average uncertainty reduction and the average realized pre-FOMC drift 24 hours before announcements in the model and in the data. The black

[39] Note that the evolution of the probability estimation is endogenously determined in the model, which affects the trend of the pre-FOMC announcement drift.

lines represent the fluctuations in the data. To examine the overnight price dynamics, I calculate the pre-FOMC drift by E-mini instead of the S&P 500 index. VIX is only allowed to trade during regular trading hours between 9:30 a.m. and 4:15 p.m. ET before April 2016.[40] There is no overnight data. Thus, I plot the implied variance reduction from 3 hours before announcements comparing to that 24 hours before FOMC announcements.[41]

The dotted red lines indicate the case under calibrated parameters. For comparison, I also show the case under the risk-neutral market makers (the dashed blue lines) and keep other parameters the same. Both cases match the uncertainty reduction pretty well. However, only under the risk-averse market makers, there is a strictly positive pre-FOMC announcement drift that is consistent with the data. Besides, the dotted red line matches the nonlinear trend of the pre-FOMC drift pretty well that the cumulative return grows faster when approaching FOMC announcements. The intuition behind this is the following. When the insider is informed, as time goes by, the cumulative order flow reveals more private information that speeds up the probability estimation of the market makers. It results in a faster uncertainty reduction, which is associated with the deeper pre-FOMC announcement drift.

I follow equation (1.33) to calculate the order imbalances in the model. Figure 1.8 plots the distribution conditional on whether the insider is informed or not. When the insider is informed, the average of the order imbalances is significantly positive due to the positive excess return before announcements. When the insider is not informed, the average is not significantly different from zero. These model implications are consistent with the empirical facts shown in Figure 1.3.

1.5 Further implications

In addition to the uncertainty reduction before announcements shown in section 1.1, I demonstrate that the asset market fluctuations around FOMC announcements are consistent with the model's predictions. First, I provide empirical support that the

[40] In April 2016, Cboe began dissemination of the VIX Index outside of U.S. trading hours so that it can be traded during "extended trading hours" between 3 a.m. and 9:15 a.m ET.

[41] I can extend the implied variance change to 5 hours before announcements if I delete the FOMC announcements that happen at 12:30 p.m. ET. The same pattern holds.

pre-FOMC announcement drift is consistent with the risk-reduction explanation before announcements from private information in my model, instead of unexpectedly good news. Second, I explain the timing of pre-FOMC drift that occurs 24 hours before announcements when private information is probably known way before. Third, I explain the two distinctive patterns to equity returns on days with an FOMC announcement since April 2011, documented by Boguth, Gregoire, and Martineau (2019). Fourth, I show that my model can reconcile with the auxiliary puzzle documented in Lucca and Moench (2015)—the absence of the pre-FOMC drift in fixed income instruments.

1.5.1 Risk-reduction explanation before FOMC announcements

My model predicts that a more substantial uncertainty reduction is associated with the stronger pre-FOMC drift. In Figure 1.2, I already show that the pre-FOMC drift only exists when there is uncertainty reduction before announcements. To more formally assess the impact of uncertainty reduction on the excess stock market returns prior to FOMC announcements, I run the following regression

$$\text{Cum. Return}_t = \alpha + \beta \Delta \text{VIX}_t + \varepsilon_t,$$

where both ΔVIX_t and Cum.Return_t are calculated from 2 p.m on pre-announcement date to announcement time windows, and t represents each FOMC announcement. As shown in Table 1.5, on average when VIX decreases 1 percent before FOMC news, the cumulative return increases 51.3 basis points.[42] In terms of the high-reduction group, since the constant term α is not significantly different from zero, the single variable uncertainty reduction can fully account for the pre-FOMC drift, which is consistent with my model.

To understand stock-bond dynamics, Cieslak and Pang (2020) decompose daily innovations in stock returns and yield changes into 4 orthogonal sources of news: growth news (cash-flow risk), monetary news (pure discount-rate risk), hedging premium news (compensation for cash-flow risk), and common premium news (compensation for discount-rate risk). They find risk-premium shocks generate 69% of the average FOMC-day increase (split into 36% and 33% contributions of the common premium and the hedging

[42] This is consistent with the simple dummy variable regression model in Table 1.6, which indicates the change of VIX before announcements itself can explain a large fraction of the pre-announcement drift.

premium, respectively).[43] This is consistent with my model's mechanism—along with the uncertainty reduction out of insider trading, the pre-FOMC drift is determined by the negative covariance between the innovation to the A-SDF (compensation for discount-rate risk) and the asset value (compensation for cash-flow risk).

The information channel I emphasize is consistent with recent work by Nakamura and Steinsson (2018).[44] They find that Federal Reserve announcements affect beliefs not only about monetary policy but also about economic fundamentals. Both of the two measures of monetary policy surprises constructed by Nakamura and Steinsson (2018) are indifferent from zero on average. Cieslak and Pang (2020) also show that the average growth news component is close to zero and conclude that the FOMC days are not associated with systematically positive or negative news about the economy. Therefore, the pre-FOMC announcement drift can not be driven by unexpectedly good news.

1.5.2 The timing of the pre-FOMC announcement drift

To fully account for the pre-FOMC announcement drift, timing is another puzzle that needs to be explained: Why does it occur 24 hours prior to announcements when private information is probably known way before?

I extend the benchmark model such that the insider knows private information earlier than 24 hours before FOMC announcements and chooses when to start to trade so that she can maximize her profits. Figure 1.1 shows the market uncertainty increases from Day -3 and does not decrease until the insider starts to trade 24 hours prior to FOMC announcements.[45] The insider's expected profit out of the asymmetric information increases in the market's uncertainty because the insider has relatively more private information when the market is nosier. Given that market uncertainty increases before FOMC announcements, the insider would like to trade later instead of trading immediately when receiving private information. However, she can not trade

[43] This number will be higher if they focus on the pre-FOMC drift instead of daily close-to-close returns since there should be no monetary policy shock before announcements.

[44] Jarocinski and Karadi (2020) find similar results—ignoring the central bank information shocks biases the inference on monetary policy nonneutrality. The effect will be stronger if they focus on the pre-FOMC period.

[45] Consistent with the empirical fact, the uncertainty keeps increasing before announcements in the model as long as the insider has not started to trade, which is captured by equation (1.5).

too late since she needs a large amount of liquidity trading to hide her information. Otherwise, the profit will be smaller. It implies that the insider faces the trade-off between uncertainty and liquidity when she decides when to start to trade. Under my current calibration, the insider optimally chooses to start to trade 24 hours before announcements to maximize the expected profits. Therefore, my paper also explains the timing of the pre-FOMC drift, which is another important feature of the pre-FOMC puzzle.

1.5.3 Two distinctive patterns to equity returns: press conferences

Since April 2011, the Chair of the FOMC has been giving a press conference at every other FOMC meeting.[46] At these meetings the FOMC also releases the summary of its members' economic projections (SEP), so that three forms of communication take place: the FOMC statement, the SEP, and the press conference with the Chair.

Boguth, Gregoire, and Martineau (2019) study the impact of the press conferences and find that the pre-FOMC drift is limited to announcements with press conferences since April 2011. Figure 1.9 shows the average cumulative return on the S&P 500 index on two-day windows with and without press conferences from April 2011.[47] The left top panel shows the VIX index with press conferences decreases significantly before announcements with an average 25.4 basis points pre-FOMC return. Besides, the VIX index keeps decreasing after FOMC announcements indicating the high uncertainty associated with press conferences. However, the return before announcements without press conferences is not significantly different from zero. Meanwhile, the VIX index almost does not change before and after announcements.

The two distinctive patterns emphasize only when the upcoming FOMC announcements are informative (i.e., with significant uncertainty resolution), there is a pre-FOMC drift.[48] It agrees with my model that the asymmetric information is less influential when the FOMC news is not informative.[49]

[46] From January 2019, the Chairman of the Federal Reserve holds a press conference after each meeting.

[47] From April 2011, some FOMC announcements happen before 2 p.m. The same pattern holds when I precisely capture the timing of all FOMC announcements.

[48] These results are consistent with findings in Boguth, Gregoire, and Martineau (2019) that announcements on days without press conferences convey less price-relevant information.

[49] This explanation agrees with Kurov, Sancetta, Strasser, and Wolfe (2019). They conclude the

1.5.4 The absence of the pre-FOMC drift in fixed income instruments

In this section, I show my model can explain the apparent lack of the pre-FOMC drift in fixed income instruments. In the model, there is a positive (negative) pre-announcement drift if the risk exposure β of the asset to the underlying fundamental is positive (negative). In other words, the sign of the pre-FOMC drift in the model depends on whether the asset is risky or a hedge.

Campbell, Sunderam, and Viceira (2017) and Campbell, Pflueger, and Viceira (2020) document that nominal Treasury bonds changed from risky (positively correlated with stocks) in the 1980s and 1990s to safe (negatively correlated with stocks) in the first decade of the 2000s. The average of time-varying betas of nominal bond is close to zero from 1996 to 2019, which results in the absence of the pre-FOMC drift in fixed income instruments. My results agree with Cieslak and Pang (2020), which find the reduction in the common premium is offset by a decline in the value of the hedging premium, making the overall bond market response economically small and statistically insignificant on FOMC days.

1.6 Discussion

Section 1.1 provides empirical support and discusses potential sources of private information prior to FOMC announcements. In this section, I discuss the challenges that the private information explanation faces in the literature. After that, I talk about other explanations.

1.6.1 Challenges of the private information explanation

Most mentioned challenges of the private information in the literature are related to consistently positive FOMC news, such as Lucca and Moench (2015), Bilyi (2018), Laarits (2019), and Cocoma (2018). However, this paper studies the resolution of uncertainty via private information results in an upward drift in market prices even if the private news is on average neutral. Therefore, their arguments do not apply to this framework. For example, they argue that if the drift is caused by private information, the realized

weaken of the average pre-FOMC drift since April 2011 comes from the lower uncertainty.

pre-announcement return should predict with a positive sign that the market response to the announcement. While my model predicts there is no correlation between the pre-FOMC returns and announcement returns, which is supported in Luuca and Moench (2015).[50]

Besides, as shown in section 1.4, the model only requires that the insider is informed for *some* FOMC announcements instead of all of them. This is consistent with Figure 1.2 that uncertainty reduction only happens before some FOMC announcements, which is associated with the pre-FOMC drift.

1.6.2 Other explanations

Hu, Pan, Wang, and Zhu (2020) and Laarits (2019) contribute the pre-FOMC drift to uncertainty reduction before FOMC news in a representative-agent framework. Though their stories are risk-based as this paper, there are two main differences: (1) The market news carries two different types of risks, and only one type of risk is resolved before FOMC announcements. (2) All investors observe the resolved information at the same time, i.e., there is no asymmetric information caused by private information.[51] Both papers face the challenge of explaining that there is no uncertainty resolution for some FOMC announcements and the pre-FOMC drift's timing. In addition to that, their models predict substantial post-announcement returns when the other risk is resolved. However, the post-FOMC announcement return is not significantly different zero, as documented in Lucca and Moench (2015).

To account for the pre-FOMC drift, Cocoma (2020) studies a model where both the risk and disagreement are very low before announcements and very high after announcements. However, the risk pattern is the opposite of Hu, Pan, Wang, and Zhu (2020), that find the risk (measured by the VIX index) starts to increase six days before announcements, then decreases from 24 hours before FOMC news until the end of days

[50] Section 24.5 of Back (2017) shows the insider trades more slowly than in the standard model since the insider considers the effect of her trades on the price. This can potentially explain why the trading volume is lower before FOMC announcements.

[51] The two risks are different in these two papers. In Hu, Pan, Wang, and Zhu (2020), the uncertainty about the potential magnitude of the news' market impact is resolved before announcements, while the risk associated with the news realization itself is resolved upon announcements. Laarits (2019) assume there are two types of announcements the Fed will make, which will reveal either monetary policy stance or long-term growth expectations. All investors learn the type of announcements before FOMC news, which resolves part of the uncertainty.

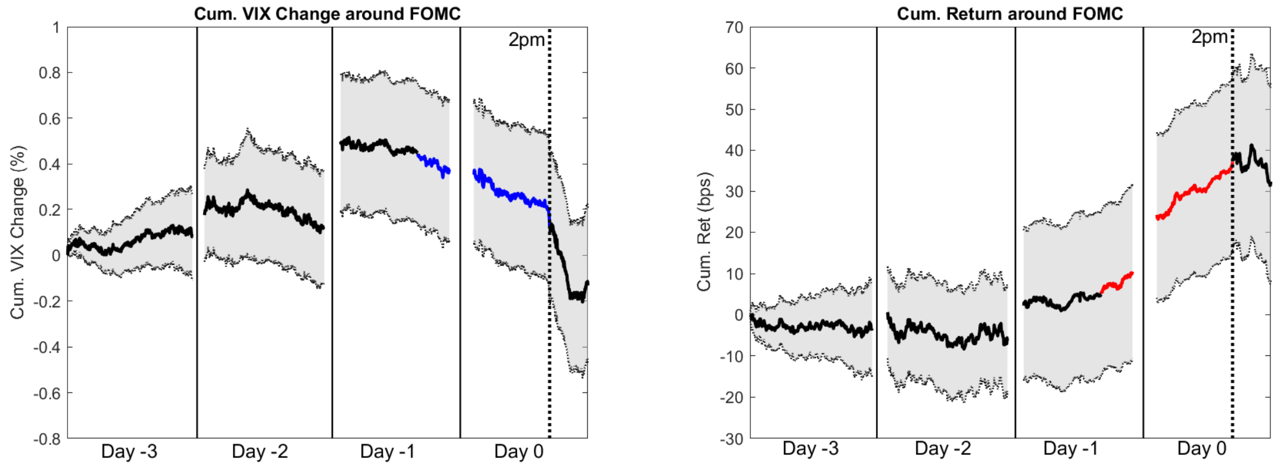
with announcements. Also, Ying (2020) shows that both call and put open interest decrease significantly at the end of days with announcements, which can not be explained by higher disagreement after FOMC news.

1.7 Conclusion

The substantial stock market return before announcements when the FOMC members refrain from discussions of monetary policy information provides a notable challenge to standard asset pricing theory. In this paper, I propose and test the private information explanation to account for the pre-FOMC announcement drift. When the uncertainty reduces before FOMC, the abnormal order imbalances are 1.85-2.17% higher in the direction of the realized return in the 24-hour window before FOMC announcements. It provides evidence consistent with informed trading when the pre-FOMC drift occurs. I integrate Kyle's (1985) model into a standard consumption-based asset pricing framework where the market makers require compensation for the risk of assets' fundamentals. Insider trading resolves uncertainty gradually and results in an upward drift in market prices, even the private news is on average neutral. The limit of the equilibrium is the traditional Kyle model as the market makers converge to be risk-neutral. The convergence result demonstrates a strictly positive pre-FOMC drift if and only if the market makers are compensated for the risk of assets' fundamentals.

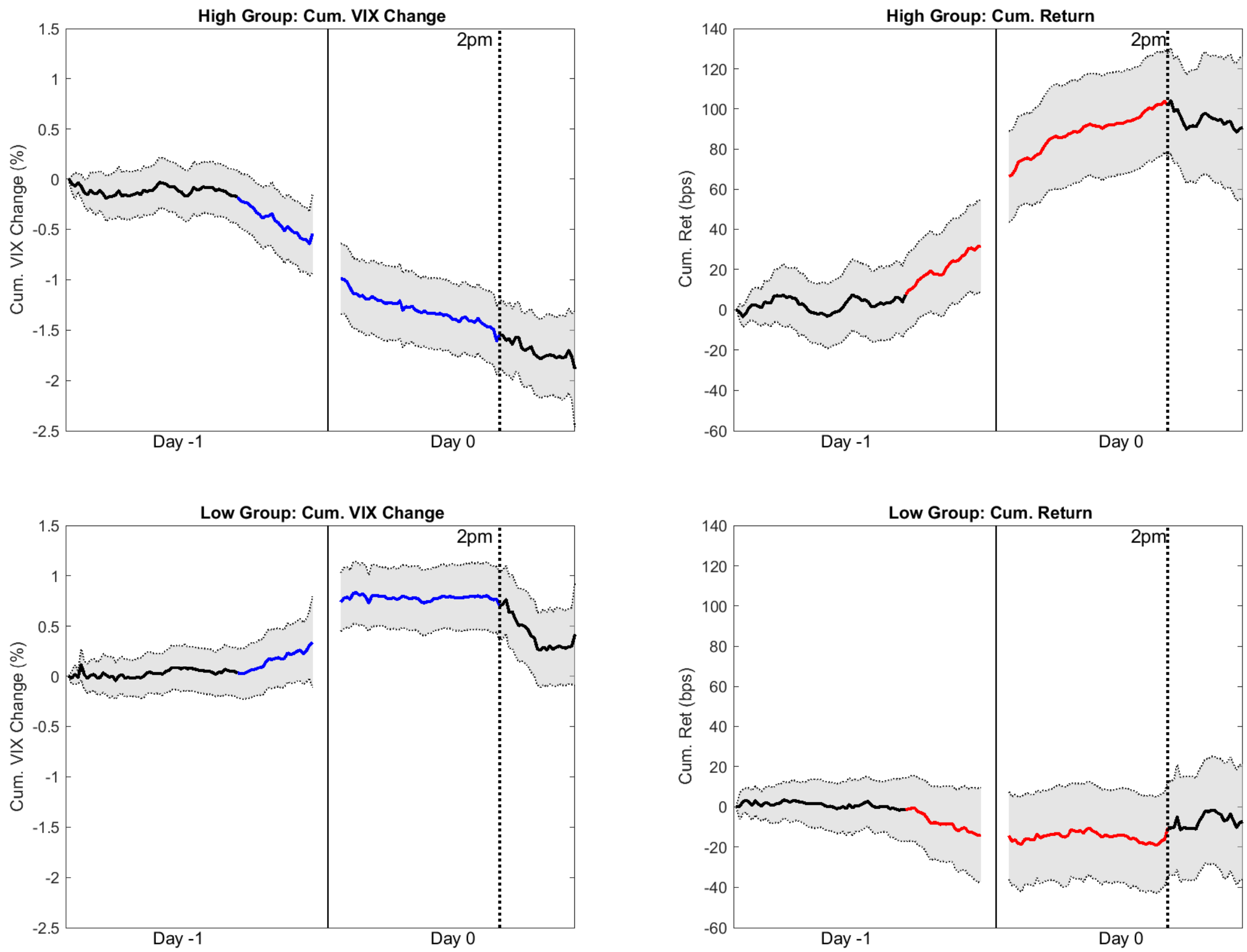
This paper provides a general framework to account for other pre-event drifts. A large group of papers treats average abnormal positive excess returns before events as evidence of insider trading and tests the market liquidity measure inspired by Kyle. While in the standard Kyle-type models, the expected average excess return before announcements is zero due to the risk-neutral market makers. Therefore, this paper provides a general theoretical framework for other pre-event drifts if the risk of the news is priced in the pricing kernel. The different equilibrium implications comparing to the standard Kyle-type models offer new insights into how private news affects asset prices, volatility, volume, and market liquidity. I leave these interesting directions for future work.

Figure 1.1: The average cumulative VIX change and return around FOMC announcements



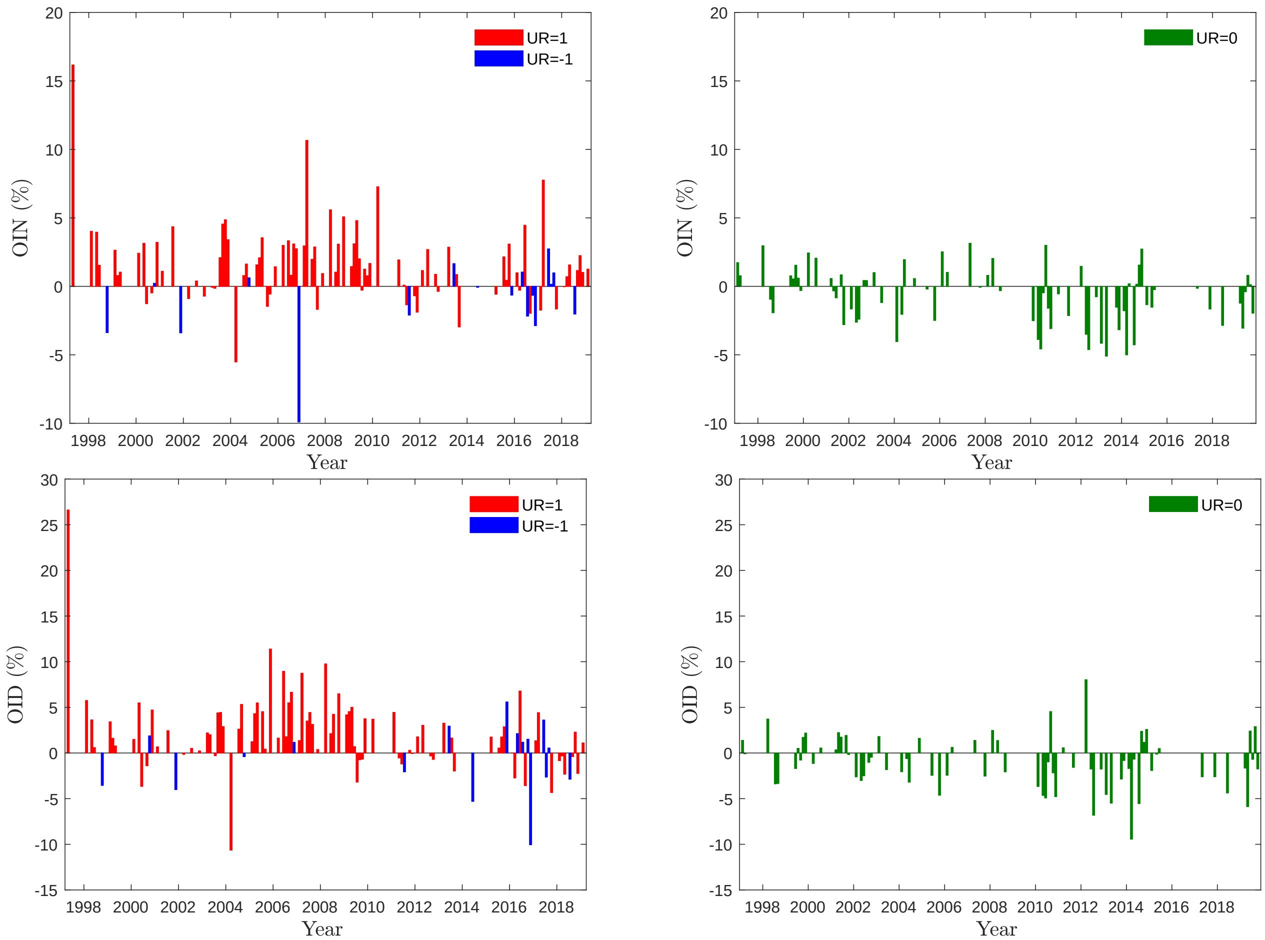
This figure shows the average cumulative VIX change and average cumulative return on the S&P 500 index on four-day windows from 1996 to 2019. The solid line of the left (right) panel is the average cumulative VIX change (average cumulative return of the SPX) from 9:30 a.m. ET on three days prior to scheduled FOMC announcements to 4:00 p.m. ET on days with scheduled FOMC announcements (labeled as Day 0). The blue (red) solid line indicates the VIX change (cumulative return of the SPX) on the 2 p.m.-to-2 p.m. pre-FOMC window. The gray shaded areas are pointwise 95% confidence bands around the average. The sample period is from January 1996 to December 2019. The dashed vertical line is set at 2:00 p.m. ET, when FOMC announcements are typically just released or 15 minutes before the release.

Figure 1.2: Classifications of FOMC meetings: sort on the reduction of uncertainty



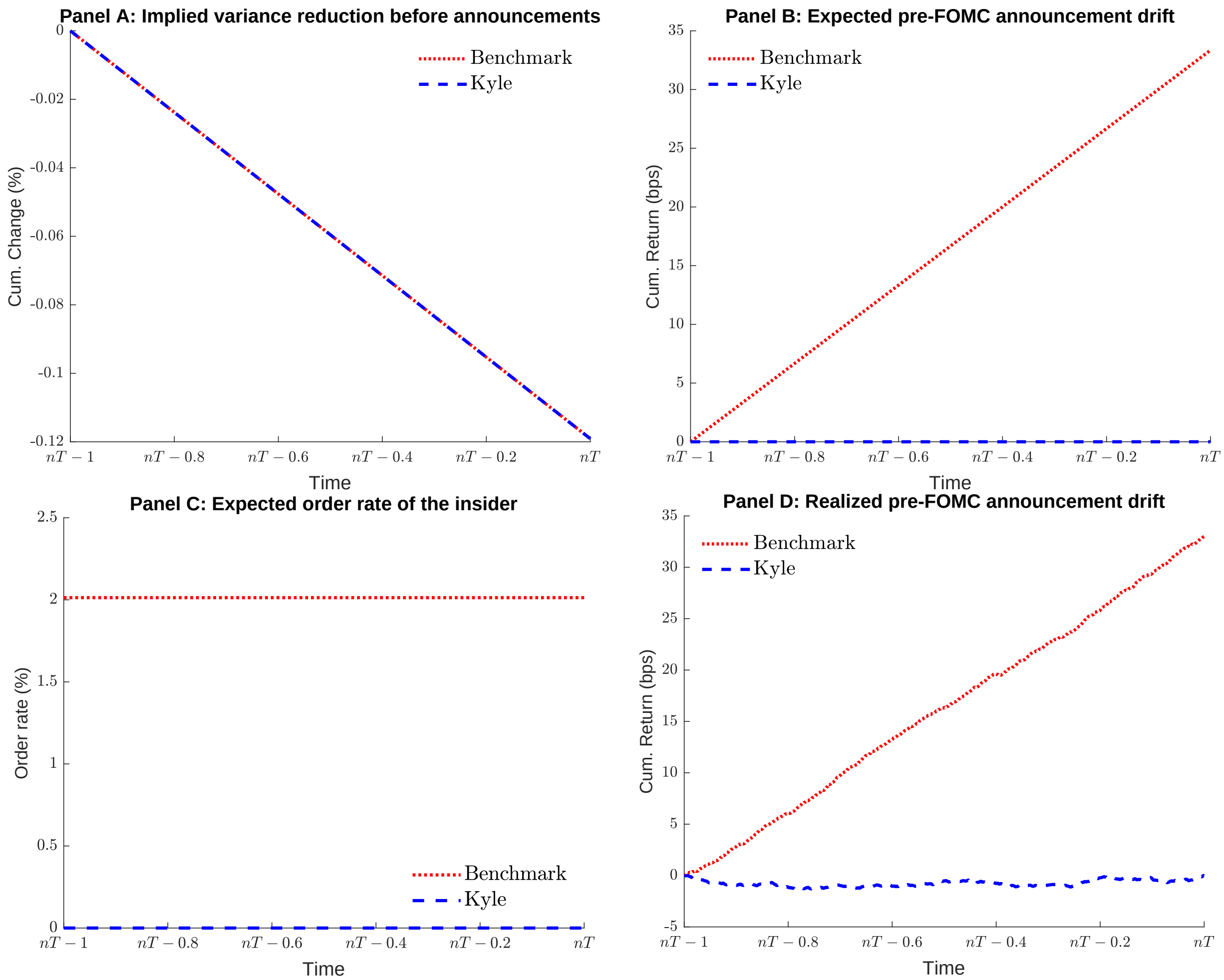
This figure shows the average cumulative return on the S&P 500 index on two-day windows for the high group and low group, respectively, where I sort the reduction of uncertainty in the 24-hour pre-FOMC window into terciles. The blue (red) solid line indicates the VIX change (cumulative return of the SPX) on the 2 p.m.-to-2 p.m. pre-FOMC window.

Figure 1.3: Measurement of informed trading: order imbalance



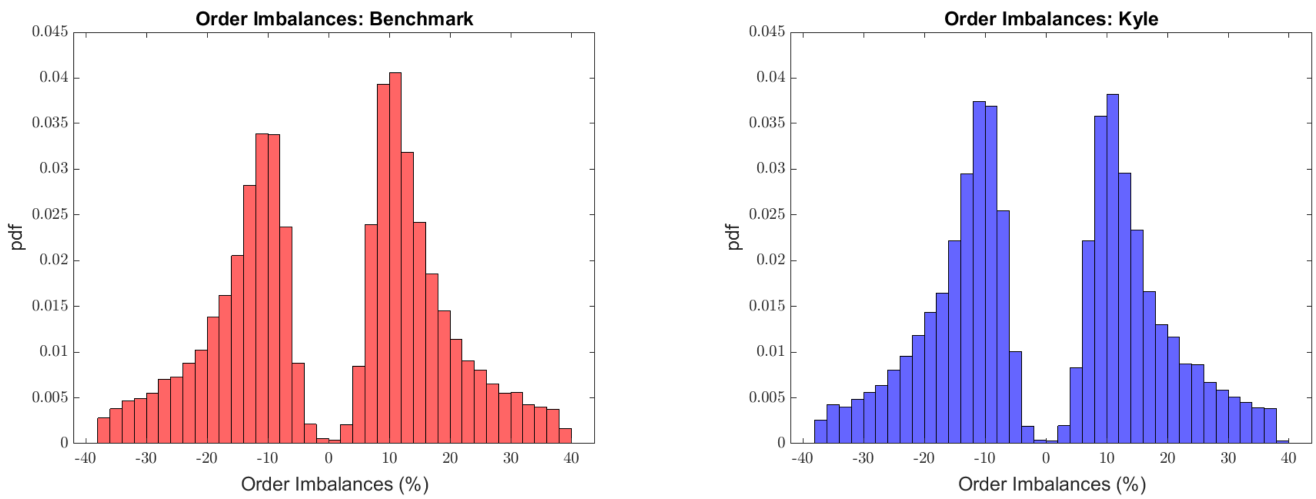
This figure shows the order imbalance based on number of trades (OIN) and dollar volume (OID) of E-mini in the 24-hour window before FOMC announcements. Red (blue) bars represent the order imbalance when the uncertainty reduces, and the cumulative return is positive (negative) over the 24-hour window, i.e., $UR = 1$ ($UR = -1$). Black bars represent the average order imbalance when the uncertainty does not reduce in the 24-hour window before FOMC announcements.

Figure 1.4: Model implications: uncertainty, expected order rate, and return



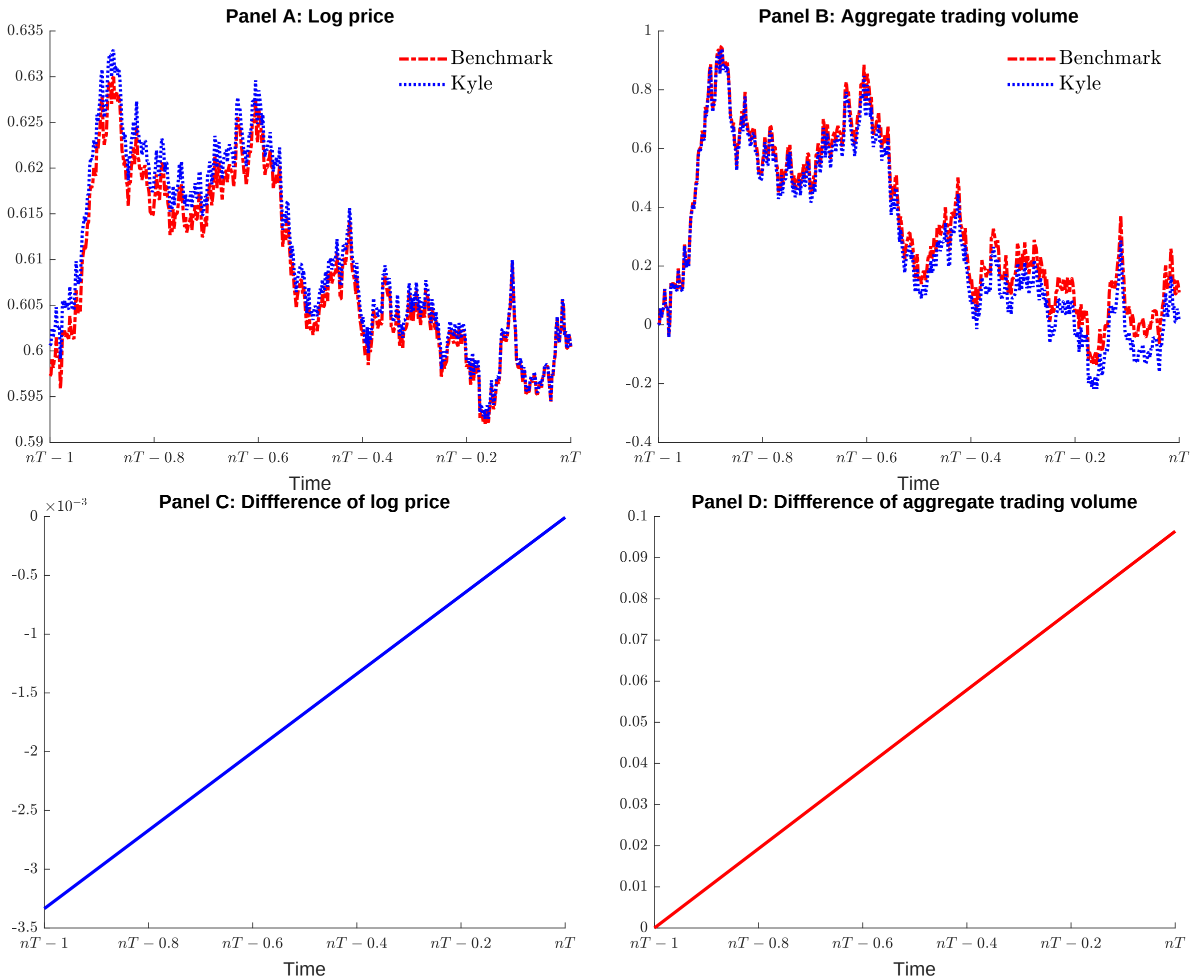
This figure shows the model implications for the cases with the risk-compensated market makers (Benchmark) and the risk-neutral market makers (Kyle) as a function of time, respectively. Panel A plots the implied variance change before announcements. Panel B plots the expected pre-FOMC announcement excess return. Panel C plots the expected insider's order rate under the market makers' filtration $\mathcal{F}_t^Y$. Panel D plots the average realized pre-FOMC announcement return, which is computed from 10,000 parallel samples. The parameters are reported in Table 1.1.

Figure 1.5: Model implications: order imbalances



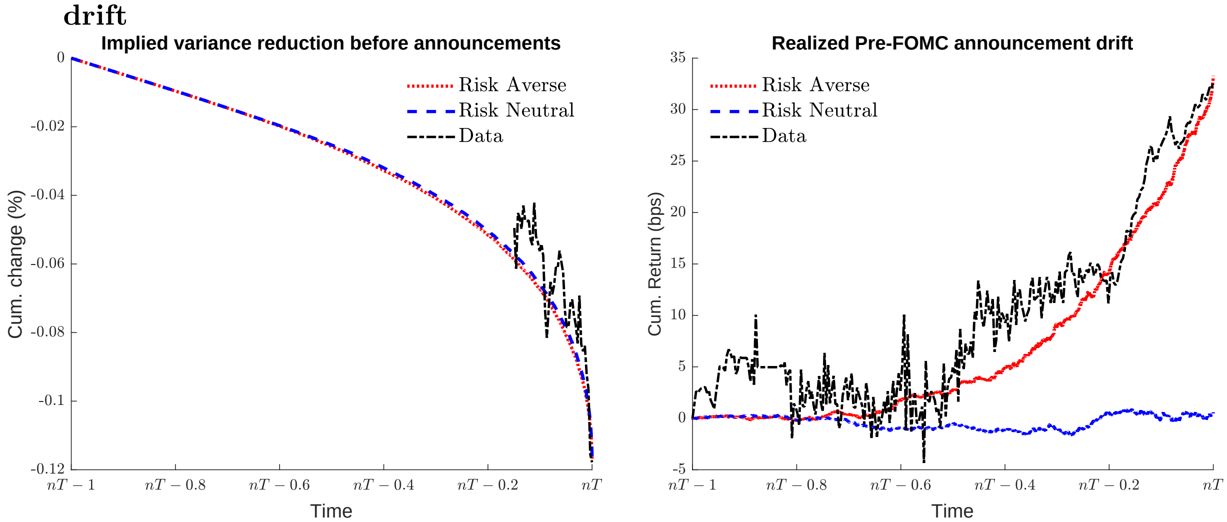
This figure shows the distribution of the order imbalances for the cases with the risk-compensated market makers (Benchmark) and the risk-neutral market makers (Kyle). I simulate 50,000 times in the frequency of one minute and winsorize the order imbalances at the 10th and 90th percentiles to mitigate the effect of outliers. The parameters are reported in Table 1.1.

Figure 1.6: Model implications: under one realized path of Z_t



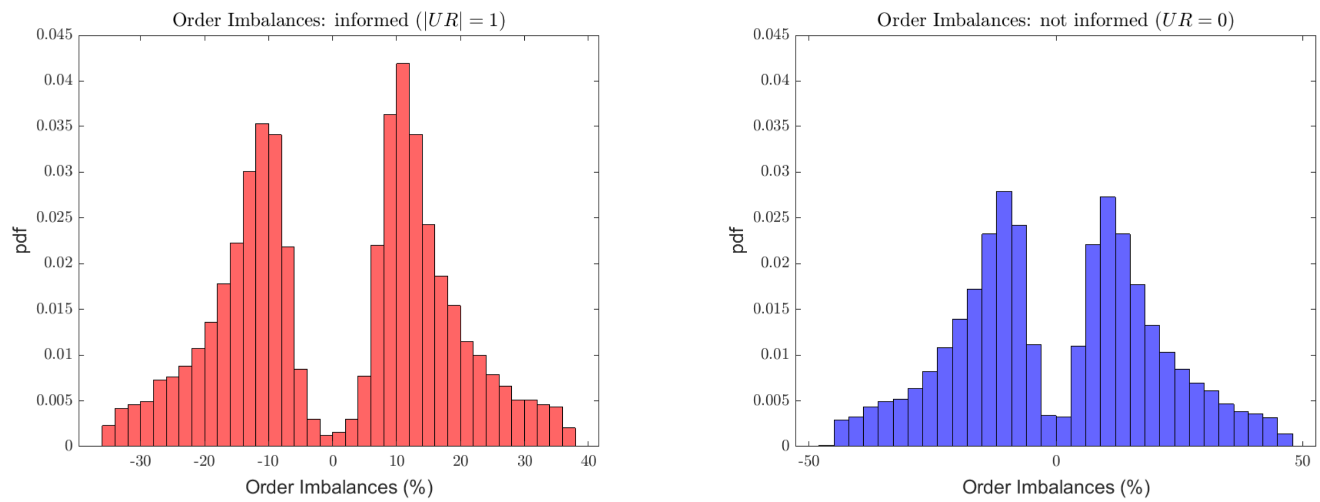
This figure shows the model implications for the cases with the risk-compensated market makers (Benchmark) and the risk-neutral market makers (Kyle) as a function of time, under one realized path of Z_t. Panel A plots the dynamics of the log price. Panel B plots the aggregate trading flow. Panel C plots the difference of log price $\log P_t - \log P_t^{\text{Kyle}}$. Panel D plots the difference of the aggregate trading flow $Y_t - Y_t^{\text{Kyle}}$. The parameters are reported in Table 1.1.

Figure 1.7: Model implications: uncertainty reduction and the pre-FOMC drift



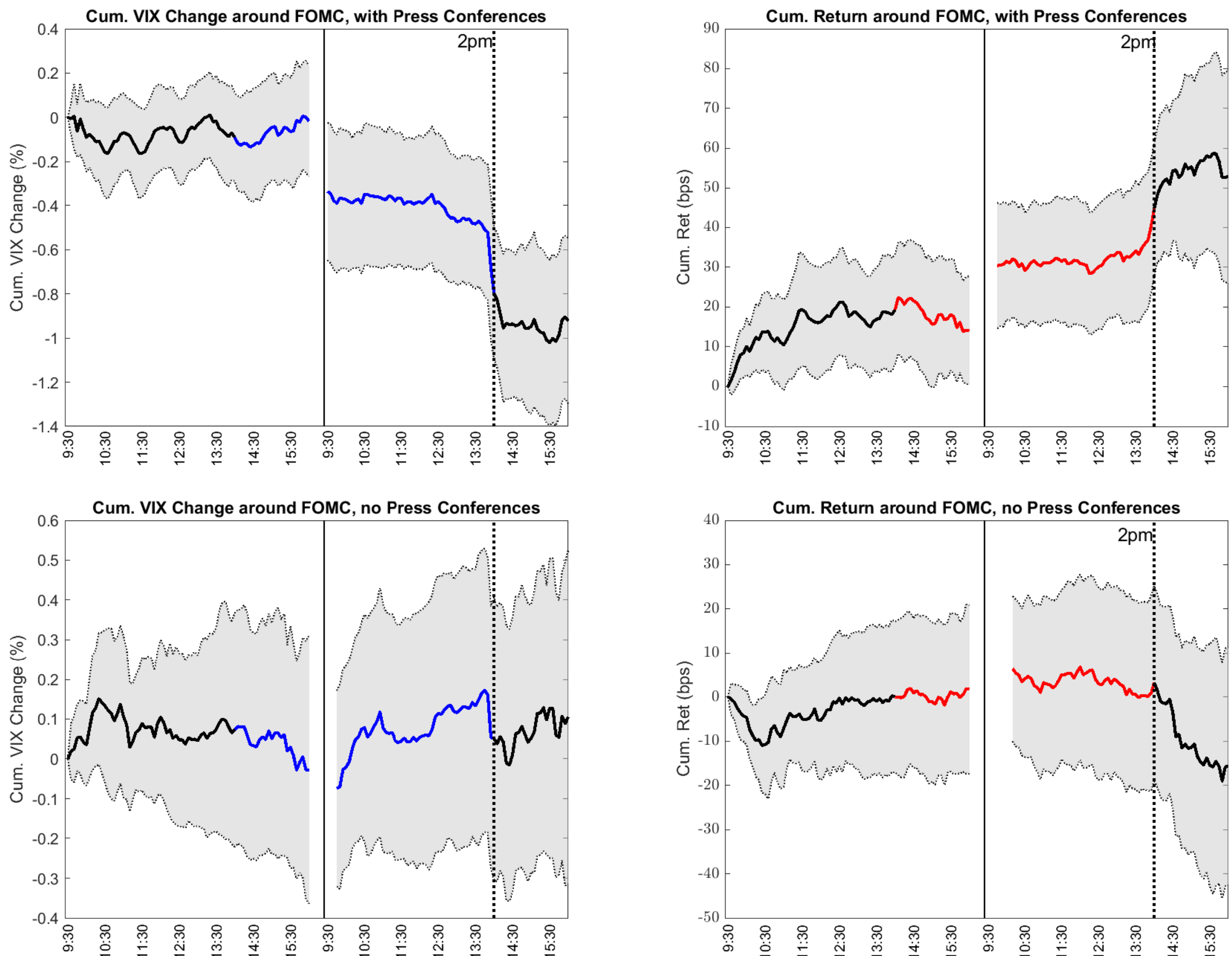
This figure plots the average uncertainty reduction and the average realized pre-FOMC drift 24 hours before announcements in the model and in the data. The dotted red lines and the dashed blue lines indicate the cases with the risk-averse market makers and the risk-neutral market makers, respectively. The black lines show the change of VIX^2 and the cumulative return of E-mini around FOMC announcements in the data. The parameters are reported in Table 1.1.

Figure 1.8: Model implications: order imbalances



This figure shows the distribution of the order imbalances in the model conditional on whether the insider is informed or not. I simulate 50,000 times in the frequency of one minute and winsorize the order imbalances at the 10th and 90th percentiles to mitigate the effect of outliers. The parameters are reported in Table 1.1.

Figure 1.9: Classifications of FOMC meetings: press conferences



This figure shows the average cumulative return on the S&P 500 index on two-day windows with and without press conferences from April 2011 to December 2019. The blue (red) solid line indicates the VIX change (cumulative return of the SPX) on the 2 p.m.-to-2 p.m. pre-FOMC window.

Table 1.1: **Parameters**

The model is calibrated at annually frequency. I assume the prescheduled announcements happen at the monthly frequency, that is, $T = \frac{1}{12}$.

Parameter	symbol	value
Aggregate output		
long run output growth rate	$\bar{m}$	1.50%
volatility of aggregate consumption	σ_C	3.16%
persistence of the AR(1) process	a_m	4.5%
volatility of the AR(1) process (benchmark)	σ_m	0.30%
volatility of the AR(1) process (full model)	σ_m	0.43%
Uncertainty and asset value		
the transparency of announcements	σ_s^2	1.6×10^{-5}
the exposure of the risky asset	β	3
Preference		
risk aversion	γ	6.6
elasticity of intertemporal substitution	ψ	1
subjective discount factor	ρ	0.005
Parameters in the full model		
prior of the probability that the insider is informed	π_{nT-1}	0.2
fraction of the informed insider across announcements	η	0.5

Table 1.2: **Summary Statistics on S&P500 Index Excess Returns and Changes in VIX**.

Note: This table reports summary statistics for the pre-announcement day 2 p.m (-1) – announcement (ann) and announcement – close changes in VIX (ΔVIX) and cumulative excess returns on the S&P500 (*Cum.Return*). The close time is 3:55 p.m. The samples are: (1) All FOMC announcements, (2 and 3) FOMC announcements sorted on uncertainty, which is first and third tertiles of changes in VIX (ΔVIX_{t-1}) between open and 2 p.m on pre-announcement dates, and (4 and 5) FOMC announcements with and without FOMC Press Conference. "Sharpe ratio" is the annualized Sharpe ratio on FOMC announcement returns. The sample period is from 1996:01 to 2019:11, and from 2011:04 for press conference sample. "No. of FOMC" is the number of FOMC in each subset. t-statistics for the mean are reported in parentheses. ***, **, and * represent 1%, 5%, and 10%.

		(1) All	Sort on Uncertainty		Press Conference	
			(2) High	(3) Low	(4) Yes	(5) No
ΔVIX (%)	2 p.m (-1)-ann	-0.300***	-1.459***	0.748***	-0.734***	-0.072
		(-3.402)	(-14.020)	(5.485)	(-5.213)	(-0.465)
	ann-close	-0.318***	-0.246*	-0.487***	-0.203*	-0.035
		(-4.405)	(-1.819)	(-3.288)	(-1.773)	(-0.181)
	No. of FOMC	187	61	63	34	35
Cum.Return (%)	2 p.m (-1)-ann	0.332***	0.944***	-0.154	0.254***	0.014
		(5.636)	(9.016)	(-1.671)	(3.314)	(0.158)
	ann-close	-0.030	-0.064	0.063	0.099	-0.157
		(-0.446)	(-0.478)	(0.603)	(1.090)	(-1.328)
	No. of FOMC	187	61	63	34	35

Table 1.3: **Compare the order imbalances conditional on FOMC announcement indicators.**

Note: This table compares the level of order imbalances of the E-mini Standard & Poor's 500 futures in the 24-hour window before FOMC announcements. OIN is the order imbalance defined as $\frac{B-S}{B+S}$, where B (S) is the aggregate buyer-initiated (seller-initiated) trading volume as measured by number of trades. OID is calculated similarly using dollar trading volume. For every FOMC announcement, I calculate the average level of order imbalances in the 24-hour window before FOMC with and without uncertainty reduction. Column (1) reports the average level of order imbalances on announcements with the pre-FOMC uncertainty reduction ($UR = \pm 1$). Column (2) reports the average level of order imbalances on announcements without the pre-FOMC uncertainty reduction ($UR = 0$). Column (3) reports the difference between columns (1) and (2). The sample period is from 1996:01 to 2019:11. t-statistics are reported in parentheses. ***, **, and * represent 1%, 5%, and 10%.

	$UR = \pm 1$	$UR = 0$	Difference
	(1)	(2)	(3)
$OIN(\%)$	1.272***	-0.719***	1.991***
	(4.22)	(-2.91)	(5.11)
$OID(\%)$	1.789***	-1.056***	2.845***
	(4.21)	(-3.13)	(5.24)

Table 1.4: **Order imbalances conditional on FOMC announcement indicators.**

Note: This table reports ordinary least squares estimates of the relation between event-time order imbalances in the E-mini Standard & Poor's 500 futures market and announcement day indicators. For each FOMC announcement, the sample includes the announcement day ($ANN = 1$) and non-announcement days in the prior 21 trading days or since the last announcement ($ANN = 0$). OIN is the order imbalance defined as $\frac{B-S}{B+S}$, where B (S) is the aggregate buyer-initiated (seller-initiated) trading volume as measured by number of trades. OID is calculated similarly using dollar trading volume. Both dependent variables are calculated in three event windows: [-24H, 0], [-24H, -12H], and [-12H, 0], where zero is the official release time of the FOMC announcement and the time unit is an hour. The uncertainty-reduced indicator, UR, is equal to one (negative one) for announcements that the pre-FOMC realized return is positive (negative) under uncertainty reduction before FOMC and zero otherwise. The sample period is from 1996:01 to 2019:11. t-statistics are reported in parentheses. ***, **, and * represent 1%, 5%, and 10%.

	[-24H, 0]		[-24H, -12H]		[-12H, 0]	
	(1) OIN	(2) OID	(3) OIN	(4) OID	(5) OIN	(6) OID
Constant	-0.047	-0.154**	0.210**	-0.065	-0.104	-0.182**
	(-0.98)	(-2.29)	(2.23)	(-0.49)	(-1.52)	(-1.98)
ANN	-0.222	-0.073	-0.562	-0.319	-0.172	-0.019
	(-0.87)	(-0.20)	(-1.11)	(-0.45)	(-0.47)	(-0.04)
UR	1.851***	2.167***	2.097***	2.213**	1.694***	2.052***
	(5.61)	(4.60)	(3.20)	(2.39)	(3.55)	(3.21)

Table 1.5: **Returns on the S&P500 Index**

Note: This table shows results for regressing the changes in VIX (ΔVIX) on the cumulative excess returns on the S&P500 (*Cum.Return*), Cum. $\text{Return}_t = \alpha + \beta\Delta\text{VIX}_t + \varepsilon_t$ where both ΔVIX_t and Cum.Return_t are calculated from 2 p.m on pre-announcement date to 2 p.m on announcement date windows, and t represents each FOMC announcement. The samples are: (1) All FOMC announcements, (2 and 3) FOMC announcements sorted on uncertainty, which is first and third tertiles of changes in VIX (ΔVIX_{t-1}) between open and 2 p.m on pre-announcement dates, and (4 and 5) FOMC announcements with and without FOMC Press Conference. The sample period is from 1996:01 to 2019:11, and from 2011:04 for press conference sample. "Obs." and "No. of FOMC" are the number of observations and amount of FOMC in each subset, respectively. t-statistics are reported in parentheses. ***, **, and * represent 1%, 5%, and 10%.

	(1) All	Sort on Uncertianty		Press Conference	
		(2) High	(3) Low	(4) Yes	(5) No
ΔVIX	-0.513***	-0.603***	-0.493***	-0.346***	-0.460***
	(-16.387)	(-5.775)	(-8.405)	(-4.646)	(-7.672)
Constant	0.170***	0.056	0.207***	-0.002	-0.021
	(4.374)	(0.319)	(2.700)	(-0.031)	(-0.390)
Obs.	187	61	63	34	35
No. of FOMC	187	61	63	34	35

Table 1.6: **S&P500 Index Return Time-Series Regressions**.

Note: This table reports results for regressions of the time-series of pre-FOMC announcement returns on various explanatory variables for the sample period 1996:01 to 2019:11. The dependent variable is a time-series of cumulative excess returns on the S&P500 from 2 p.m on days before announcement to 2 p.m on days of scheduled FOMC announcements. The first independent variable in Column (1) and (2) is Pre-FOMC dummy (D_{FOMC}), which is equal to one when a scheduled FOMC announcement has been released in the following 24-hour interval and zero otherwise. The second independent variable in Column (2) is the interaction of changes in VIX and Pre-FOMC dummy ($\Delta\text{VIX} \times D_{\text{FOMC}}$). "Sharpe ratio" is the annualized Sharpe ratio on FOMC announcement returns. "Obs." and "No. of FOMC" are the number of observations and amount of FOMC in each subset, respectively. t-statistics are reported in parentheses. ***, **, and * represent 1%, 5%, and 10%.

	(1)	(2)
D_{FOMC}	0.314***	0.160**
	(3.737)	(1.857)
$\Delta\text{VIX} \times D_{\text{FOMC}}$		-0.513***
		(-7.507)
Constant	0.010	0.010
	(0.671)	(0.674)
Sharpe Ratio	1.14	1.14
Obs.	5899	5899
No. of FOMC	187	187

Chapter 2

A Model of Market Discipline

2.1 Introduction

Crucial to the goal of corporate value maximization is the disciplinary role of markets on aligning investors' and managers' incentives. Even though infrequently tapped for capital, markets invisibly guide management's use of resources, as serially poor judgment eventually necessitates costly refinance. The degree to which this channel operates, however, is an open question. How strong is it and how does it manifest itself? More broadly, how do agency and financial frictions jointly affect firm behavior and can they be altered? In spite of the importance of these questions concerning the efficacy of markets, a framework suitable for studying them has proved challenging.

In this paper, we attempt to fill this gap. We do so by developing a quantitative model that combines a dynamic agency problem with internal and costly external finance.[1] In the model, investors anticipate refinancing's effect on management's incentives and thus judiciously choose it given the firm's history. Managers understand this and, therefore, it affects their behavior today and subsequent real outcomes, which in turn feeds back into investors' expectations. We call this equilibrium effect *market discipline*.

Our paper makes two major contributions. First, by formalizing the interaction

[1] If only costly financing were specified, markets would have no scope to affect management's already optimal behavior. If only an agency conflict were present, instances of costly refinancing would be left indeterminate and remain non-measurable. The ability to finance internally makes this costly event meaningful.

between financial and agency frictions, we provide a unified framework for assessing the role of markets in shaping cash holdings, investment, payouts, compensation, and whether to refinance a firm or let it fail. Second, we derive a general formula that connects the relative sizes of agency conflicts and financial frictions to the relative allocation of free cash flow across investors and managers:

$$\frac{\text{Size of Agency Conflict}}{\text{Size of Financial Friction}} \propto \frac{\text{Marginal Cost of Delaying Payouts to Investors}}{\text{Marginal Cost of Delaying Payments to Managers}}.$$

For intuition, consider a firm with scarce cash holdings. Here the marginal cost of delaying payouts to investors is low as the likelihood of costly refinancing is large, and preferably avoided altogether. Any payments made to managers, during which agency conflicts are muted, *must* then correspond to a financial distortion. The idea is akin to using a supply shock to identify a demand elasticity and analogously implies that agency conflicts can be inferred using payouts. Thus, we show that the scale at which managers or investors are differentially paid are informative about the relative magnitude of these underlying frictions.

It is useful at this point to take a step back into the model setup of Section 2.2 to better understand our headline results. To separate interests and motivate their desire to grow the firm beyond the optimal size, we allow managers the possibility of consuming private benefits that are increasing in firm assets and resources (Jensen (1986)). Their private consumption comes at the expense of their effort to increase asset efficiency and enhance the probability of firm survival. Investors therefore write a contract that continually provides managers with the incentive to choose the appropriate action.

The contract is history-dependent and acts as the bridge through which managers internalize investors forward-looking rational expectations. The dynamics of managerial incentives, in turn, influence several firm policies including refinancing, and not just the nearest refinancing event but over all those that potentially could occur afterwards.

Our model's solution described in Section 2.3 takes the form of a partial differential equation summarized by two state variables, each normalized by assets: cash holdings and managers' stake, which measures their effective ownership of the firm and proxies for both compensation and the (inverse) agency problem in this dynamic environment. The model endogenizes four decision thresholds that collectively provide lower and upper bounds for each of these states: (1) when investors collect payouts out of cash versus (2) when the firm is refinanced; and (3) when management receives payments versus (4) when the firm is liquidated. A key challenge that we overcome is thus determining the value of the solution jointly with the shape of the state space.

Firms strive to reach the state where positive free cash flow is freely paid to both investors and managers; that is, the upper thresholds (1) and (3) apply. At this joint upper boundary, the general formula holds explicitly and with equality in our setup. It equates the allocation of free cash flow across players to the ratio of two lingering marginal costs: a tax penalty to holding cash in the firm and managers' relative impatience to investors. This ratio also defines the linear span of a second-best frontier that bounds from above and is tangent to the endogenous state space. Novelly, we quantify the size of agency conflicts and financial frictions as proportional to the distance between the upper thresholds and the frontier.

Next, moving away from the joint upper boundary (where (1) and (3) hold) to along one upper boundary (where (1) or (3) holds) implies one distortion will remain minimal while the other is aggravated. The boundary's shape is therefore informative on the relative frictions that firms currently face and naturally extends to relative allocation of free cash flow, thus closing the step in understanding the origins of the general formula.

Our measurement approach has advantages over those obtainable in classic agency or financial friction models. First, classic models measure frictions as deviations from first-best, which may be neither attainable nor serve as a reasonable benchmark. Second, they often make comparisons implicitly based on market values that are influenced by hard-to-measure discount rates. By contrast, we measure relative to the second-best frontier and we do so with quantities of cash flows that are readily observable. That payouts and payments can be observed with simple accounting data is an appealing feature with empirical potential, but is not within this paper's scope to pursue.

More broadly, a burgeoning literature studies the impact of financing frictions on the economy. It is unclear that in developed economies, however, that these are the utmost concern since capital markets are quite deep. Instead, we argue that agency frictions are likely to be more onerous. Indeed, our calibrated model of Section 2.4 suggests that they are nearly 10 times more severe than financial frictions in the United States. This is because while cash can simply be accumulated to minimize financing frictions, it is double-edged as it exacerbates the alignment of managerial incentives.

Part of our calibration is done internal to the model and targets moments in the data that speak to the rich features of market discipline—payouts to investors, managerial compensation, and frequencies of refinancing and liquidation, among others—outcomes that are obviously important to corporate finance and firm value maximization but have been little, if at all, studied together in modern structural models.

We then proceed to evaluating model counterfactuals in Section 2.5 by conducting steady state analysis in the spirit of Hopenhayn (1992) by examining how the stationary distribution shifts in response to a change in parameter. The stationary distribution encodes all of the information about the stochastic environment and policy functions of the model solution and is therefore an ideal object to study.

Among other analysis, we examine a policy counterfactual where we lower the corporate tax rate from 30 to 21 percent and raise the external cost of finance from 50 to 150 basis points, which could be implemented with a small tax on the event of refinancing. In effect, lower corporate taxes help offset the cost that investors would otherwise bear for a refinancing tax and further allows them to allocate additional cash flow to mitigating agency conflicts. While subject to caveats that we discuss in the paper, we find that its stationary equilibrium *mimics* an economy where relative agency frictions are reduced by third.

Literature

Our paper analyzes the capital market implications of dynamic agency.[2] As in DeMarzo, Fishman, He, and Wang (2012), we analyze investment in the context of a dynamic agency model. In their paper, the optimal contract relaxes agents' incentive constraint following a history of good shocks, which raises the marginal benefit of investing in more capital. However, in this paper the distinction between internal and external sources of finance are left unexplored, therefore ignoring the ex ante effect that discrete instances of refinancing have on agents' incentives.

Zwiebel (1996) critiques that a recurring, and problematic, feature of traditional agency models is that a "discipliner" is present ex ante yet absent ex post. Often in these models the discipliner sets constraints (for example, debt) that ex ante restrict

[2] A partial list is Albuquerque and Hopenhayn (2004), Quadrini (2004), Dow, Gorton, and Krishnamurthy (2005), Clementi and Hopenhayn (2006), DeMarzo and Fishman (2007), Biais, Mariotti, Plantin, and Rochet (2007), Ai and Li (2015), Morellec, Nikolov, and Schürhoff (forthcoming), Boualam (2019), Ward (2019), and Tong and Ying (2019).

managers' future decisions. If instead the discipliner were present ex post, management could still be restricted even though constraints were never set. He argues that the correct formulation of constraints, whether ex ante or ex post, is dynamically consistent, as they are in our model. Our contribution here is that our paper studies a broader range of corporate policies in the context of quantitative model.

Hartman-Glaser, Mayer, and Milbradt (2019) study moral hazard's effect in an environment where the firm accumulates cash, similar to ours. They show that when cash holdings are low, firms transfer cash flow risk to managers, hoping to minimize their desire to divert cash. In addition, they show that permitting the payments of small, negative wages to managers allows them to solve the model as a function of only one state variable. In contrast, we solve the model with two states without resorting to a restricted problem. Another novelty of our paper is that the event of refinancing itself is endogenous, a key feature in isolating the market's disciplinary effect on management's behavior.

Our theory complements the literature on financing constraints.[3] Bolton, Chen, and Wang (2011a) show that the marginal value of cash affects investment, external financing, and risk management. Specifically they show that cash holdings follow a fixed double barrier policy. The lower bound has the firm either refinancing or liquidating, depending on the choice of parameters. In contrast, our double barriers are dependent on the level of management's incentives, allowing us to study the interaction between financial and agency frictions. In addition, the important decision whether to let the firm refinance or fail is endogenous to our model.

Our paper contributes to the literature on misallocation and the measurement of distortions.[4] Pioneered by Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), a large part of the literature has focused its determinants such as financial frictions (for example, Moll (2014), Midrigan and Xu (2014), and Gopinath, Kalemli-Ozcan, Karabarbounis, and Villegas-Sanchez (2017)). Our model, in contrast, jointly determines financial frictions with agency conflicts, a feature largely ignored in the

[3] A short list is Gomes (2001), Whited and Wu (2006), Hennessey and Whited (2007), Riddick and Whited (2009), Rampini and Viswanathan (2010), Nikolov and Whited (2014), Milbradt and Oehmke (2015), and Belo, Lin, and Yang (2018).

[4] An incomplete list is King and Levine (1993), Rajan and Zingales (1998), Buera, Kaboski, and Shin (2011), Asker, Collard-Wexler, and de Loecker (2014), Haltiwanger, Kulick, and Syverson (2018).

literature. We thus provide a new rationale that possibly contributes to the mysterious decline in US allocative efficiency (Bils, Klenow, and Ruane (2020)).

2.2 Model

Here we present the model's setup. We first describe the firm's technology and how managerial effort affects its efficiency. We then introduce the flow equation for resources and costs of external financing. We finally discuss the agency problem and close the setup within the contracting environment.

2.2.1 Technology, Free Cash Flow, and Management's Effort

Capital, K, is used to produce output and evolves according to the standard accumulation equation

$$dK_t = (I_t - \delta K_t)dt, \tag{2.1}$$

where I is gross investment and $\delta \geq 0$ is the rate of depreciation. Following the literature on q theory (Hayashi (1982) and Abel and Eberly (1994)), investment incurs adjustment costs $G(I, K)$, allowing us to write the total cost of investment as $I + G(I, K)$.

Free cash flow is the cash flow available to distribute to investors after paying taxes at rate τ_Y and paying for the investment and expenses required to maintain the firm's existing operations. After optimally choosing and compensating freely adjustable labor (non-management), free cash flow is determined by a constant returns to scale technology:[5]

$$dY_t = (1 - \tau_Y)dA_t K_t - I_t dt - G(I_t, K_t)dt. \tag{2.2}$$

Asset productivity is determined by management's unobservable effort, $e_t \in \{0, 1\}$:

$$dA_t = e_t \mu dt + \sigma dZ_t, \tag{2.3}$$

[5] For a given capital stock and with freely adjustable labor, L, the firm solves the static problem $\max_L (1 - \tau_Y)(a_t K_t^{\alpha} L_t^{1-\alpha} - w_L L_t)$, where a_t is a productivity shock and w_L is the wage rate which could be stochastic. The optimal labor choice will be proportional to capital. The productivity shock dA_t used elsewhere thus depends on a_t, w_L, and α.

where dA is a productivity shock with drift $e_t\mu \geq 0$ that varies with a standard Brownian increment dZ scaled by volatility $\sigma > 0$. Effort ($e_t = 1$) enhances profits and the likelihood of firm survival. However, because only dA will be observable and contractible to investors, the extent to which management can hide their effort choice and shirk ($e_t = 0$) will scale with σ.

Even though the economic environment is *iid*, policies dictating the firm's experience under the optimal contract will be path dependent. Notably, the firm's cash flow history, which is in part determined by managerial effort, will affect equilibrium outcomes such as refinancing.

2.2.2 Internal Resources, Distributions, and Costly Refinancing

Donaldson (1984) describes a firm's resources as "the aggregate purchasing power available to management for strategic purposes during any given planning period." They are thus not limited to pecuniary things and could reflect the sophistication of business networks or even the operational efficiency of the firm—any of these could be squandered by managers. To empirically construct a law of motion, however, we restrict them to be only cash holdings.

Cash held at time t is denoted by C_t and the flow equation is

$$dC_t = dY_t + \tau_Y \delta K_t dt + r(1 - \tau_C)C_t dt + dF_t - dD_t. \tag{2.4}$$

Holdings increase with the firm's free cash flow, dY, and depreciation tax shield, $\tau_Y \delta K$. The risk-free rate of return the firm earns on its uninvested resources, $r(1-\tau_C)$, reflects a penalty on cash holdings. Funds can additionally be acquired through external financing or distributed to investors at any time.

Let F_t and D_t denote the cumulative (nondecreasing) funds acquired and dispensed by the firm up to time t and dF_t and dD_t as the respective incremental changes in these policies over the time interval $(t, t + dt)$. When financing externally and receiving funds from financial markets, firms face explicit underwriting costs and implicit costs, as investors naturally question managers' intended use of funds and the potential change to their incentives.

Modeling these costs are complicated but to provide an environment in which we can calibrate a model we follow Gomes (2001) who summarizes the costs of external financing with a fixed cost Φ and a marginal cost ϕ.[6] Together these costs imply that firms will only intermittently tap markets for funds and, when they do, raise a finite amount. To ensure firms do not outgrow financing costs, we follow Bolton, Chen, and Wang (2011a) and assume both costs scale with capital as informational or incentive costs or the effects of dilution are likely to be proportional to firm size. We denote the cumulative costs of external financing up until time t by X_t and its incremental change as dX_t.

Because financing costs scale with capital, our model provides a better approximation to the behavior of large firms, as information asymmetries between insiders and outsiders are likely greater and vary more among small firms that often have shorter track records. Moreover, we calibrate issuance costs to equity and not debt markets, as equity issuances are more likely to be informationally sensitive. Homogeneity, importantly, makes the model tractible and Eberly, Rebelo, and Vincent (2009) provide empirical support for it among large firms in Compustat.

Refinancing is not the only decision investors can make for they can also choose to let the firm fail and liquidate it. To distinguish these cases, we now turn to describing the contract between investors and managers.

2.2.3 Investors' and Management's Contract and Termination

Investors hire a management team to run the firm and write a contract that can be terminated at any time. Investors have unlimited wealth and are risk neutral and therefore discount at the risk-free rate r. Managers are also risk neutral but discount at rate $\gamma > r$.[7] They have no initial wealth and limited liability so investors cannot pay negative wages to them. At the termination time τ investors receive a fraction of assets and cash: $0 < l_K, l_C \leq 1$ of capital and resources, altogether recovering $l_K K_\tau + l_C C_\tau$.[8]

Managers receive their outside option, normalized to zero.

[6] We do not solve for the optimal external financing policy jointly with the optimal incentive contract. Rather, incentives are made compatible given this particular institutional structure of financing costs.

[7] This traditional assumption captures either their assumed impatience or in reduced-form the presence of outside investment opportunities available to them. While $\gamma = r$ may be a more neutral assumption, DeMarzo and Sannikov (2006) argue a contract can be made more robust by having investors assume that γ is higher than managers' true γ.

[8] Losing a fraction of resources is consistent losses in bankruptcy. In the broader Donaldsonian interpretation of resources, losses could reflect managers' network or specific knowledge of the inner workings of the firm. Shue (2013) documents the importance of executive peers within a MBA cohort

When management exerts no effort ($e_t = 0$) they enjoy private benefits at rate $\Lambda(K_t, C_t)\,dt$. Managers deriving benefits from capital, $\Lambda(K, \cdot)$, agrees with the long literature on empire building. There are several reasons for why they would also be expected to grow in resources, $\Lambda(\cdot, C)$. First, cash may provide funds for managers to invest in projects which offer private benefits but do not contribute to shareholder value. Because effort improves the return of productive assets ($\mathbb{E}_t[dA_t] = e_t \mu dt$) , no effort here can be viewed as putting effort into valueless projects that only managers enjoy. Second, more often than not, a cash-rich company runs the risk of being prodigal. And finally, large cash holdings remove some pressure on management to perform. Ultimately, that this function increases in both arguments is consistent with the thesis of Jensen (1986) and the international evidence on cash holdings and agency conflicts in Dittmar, Mahrt-Smith, and Servaes (2003).

We assume the capital stock K_t, cash C_t, and cumulative free cash flow Y_t are observable and contractible. From (2.1) and (2.2), investment I_t and cumulative productivity A_t can therefore be contracted upon. Investors maximize firm value by offering a contract that specifies investment, refinancing, and payout policies, $\{I\}$, $\{F\}$, and $\{D\}$, management's cumulative payments, $\{U\}$, and a termination (stopping) time, τ, all of which depend on the entire history of productivity A_t. Limited liability requires U to be nondecreasing. We let $\mathcal{C} = (I, F, D, U, \tau)$ represent the contract.

Given the contract, management chooses an effort process $\{e_t \in \{0, 1\} : 0 \leq t < \tau\}$ to solve

$$W(\mathcal{C}) = \max_{\{e_t \in \{0,1\}: 0 \leq t < \tau\}} \mathbb{E}^e \left[\int_0^\tau e^{-\gamma t} \left(dU_t + \Lambda(K_t, C_t)(1 - e_t) dt \right) \right], \tag{2.5}$$

where $\mathbb{E}^e[\cdot]$ is the expectation operator under the probability measure induced by their effort choices. Their expected utility is composed of the present discounted value of compensation and private benefits only when taking action $e_t = 0$.

At the time the contract is initiated, the firm has K_0 units of capital and C_0 units of cash. Given an initial payoff W_0 to managers, the problem investors face is

$$\begin{aligned} P(K_0, C_0, W_0) = \max_{\mathcal{C}} \mathbb{E} \left[\int_0^\tau e^{-rt} (dD_t - dF_t - dX_t - dU_t) + e^{-r\tau} \left(l_K K_\tau + l_C C_\tau \right) \right] \\ \text{s.t. } \mathcal{C} \text{ is incentive compatible and } W(\mathcal{C}) = W_0. \end{aligned} \tag{2.6}$$

in explaining firm policies that do not necessarily contribute to firm productivity.

The value to investors is the expected present discounted value of payouts, dD, less funds injected dF at cost dX and payments to managers dU, plus what they recover in liquidation.[9]

Management's payoff W_0 is determined by their relative bargaining power (DeMarzo and Sannikov (2006)). If managers possess all power, then $W_0^M \equiv \max\{W : P(K_0, C_0, W) \geq 0\}$. If however investors have all power, $W_0^I \equiv \operatorname{argmax}_{W \geq 0} P(K_0, C_0, W)$. More generally, we blend the two extremes with a parameter $\psi \in (0, 1)$ by setting $W_0 = \psi W_0^M + (1 - \psi) W_0^I$.

Incentive Compatible Contract

We focus on the case where the contract is incentive compatible and implements the efficient action $e_t = 1$ for all t. Given this contract and history up until time t, management's continuation payoff is given by

$$W_t(\mathcal{C}) = \mathbb{E}_t \left[\int_t^\tau e^{-\gamma(s-t)} dU_s \right]. \tag{2.7}$$

Note the collapse of the expectation operator $\mathbb{E}^e[\cdot]$ to the one that agrees with investors' expectation $\mathbb{E}[\cdot]$. Rational expectations on behalf of both parties has management internalizing investors' expectations and investors offering a contract consistent with those expectations. This model feature makes it dynamically consistent and formalizes the notion of market discipline affecting management's behavior ex ante.

Standard dynamic contracting theory decomposes management's incremental total compensation at time t into incremental payments, dU_t, and incremental continuation payoff, dW_t (Spear and Srivastava (1987) and Sannikov (2008)). The optimal contract compensates management for their time preference on average; the analogous promise keeping condition is $\mathbb{E}_t [dW_t + dU_t] = \gamma W_t dt$. Furthermore to maintain incentive compatibility, management's compensation must remain sufficiently sensitive to the firm's free cash flow, dY_t. Following DeMarzo and Sannikov (2006), we formulate this sensitivity, β_t, with the martingale representation theorem:

$$dW_t + dU_t = \gamma W_t dt + \beta_t \left(dY_t - \mathbb{E}_t [dY_t] \right) = \gamma W_t dt + \beta_t (1 - \tau_Y) \sigma K_t dZ_t. \tag{2.8}$$

[9] Notice that payments to managers dU are not subtracted from dC in (2.4). We thoroughly discuss this alternative setup in Appendix B. In our setup, we are implicitly assuming that firms do not refinance to pay managers current payments, like bonuses.

Agents who deviate reduce their compensation by $\beta_t(1-\tau_Y)\mu K_t dt$ and receive private benefits $\Lambda(K_t, C_t)\, dt$. Incentive compatibility is thus implemented with $\beta_t(1-\tau_Y)\mu K_t \geq \Lambda(K_t, C_t)$. Because liquidation is ex post inefficient and therefore costly to enforce, the optimal contract minimizes the likelihood of this event and sets

$$\beta_t = \frac{\Lambda(K_t, C_t)}{(1-\tau_Y)\mu K_t} \text{ for all } t. \tag{2.9}$$

Intuitively, the optimal sensitivity is a ratio of private benefits to capital's expected return potential and provides a nexus among free cash flow, compensation, capital, and cash holdings.

An important assumption is that the shocks to agents' continuation utility are all local; that is, for a given C, K, and W, the instantaneous forward distribution of 2.8 across two different economies are identical. Of course, firm value is generally not, and therefore dynamics will in general differ. This assumption is how we choose to model market discipline as an indirect force that invisibly guides managers' behavior.

Given the functional form of $\Lambda(\cdot)$ it naturally implies that for a given beneficial shock to free cash flow ($dZ_t > 0$), greater capital or resources are associated with larger increases in managers' compensation, consistent with Edmans, Gabaix, and Landier (2008). In contrast, a series of negative shocks will result in either refinancing or else the termination of the contract and liquidation of the firm. For example if incentives have become too poor, investors will decide not to refinance the firm and let it fail. We now formalize this important decision while discussing the solution to our model.

2.3 Model Solution

Here we describe some properties of the solution to (2.6). Management's continuation payoff W_t in (2.7) is a state variable that summarizes management's current incentives that reflect their expected path of compensation and the likelihood of contract termination. Capital K_t captures the history of investment via (2.1). The firm's cash holdings C_t track the histories of refinancing and payouts. Altogether, whatever the history of the firm up until date t, the only relevant state variables are K_t, C_t, and W_t and, therefore, investors' value function at time t, $P(K_t, C_t, W_t)$, can be solved with a Hamilton-Jacobi-Bellman (HJB) equation.

We assume that adjustment costs, $G(I, K)$, and private benefits $\Lambda(K, C)$ are homogeneous of degree one in their arguments. The total cost of investment is thus $I + G(I, K) = Kg(i)$, where $i = I/K$ is the investment rate, and we specify private benefits to take a linear form

$$\Lambda(K, C) = K\lambda(c) = K\left(\lambda_K + \lambda_C c\right) \tag{2.10}$$

that separates the agency friction attributed to capital (λ_K) and resources (λ_C). Homogeneity allows us to reduce the problem to two endogenous state variables—managers' stake (their scaled continuation payoff), $w = W/K$, and scaled cash holdings, $c = C/K$—and write $p(c, w) = P(K, C, W)/K$.

Common to risk neutral models, investors optimally choose investment to equate expected returns to their required rate of return, the risk-free rate:

$$rdt = \max_i \mathbb{E}_t \left[\frac{d(Kp(c, w))}{Kp(c, w)} \right], \tag{2.11}$$

subject to the incentive compatibility constraint $\beta \geq \lambda(c)/((1 - \tau_Y)\mu)$ from (2.9) and (2.10). This equation's solution is jointly determined with the boundaries that, when present, determine refinancing, payouts to investors, payments to managers, and contract termination.

In what follows we first present the solutions to models nested by our complete model as they are simpler yet still informative. We begin with the first-best model, then the model without an agency conflict followed by one without costly refinancing, before building to our complete model.

2.3.1 First-Best Solution

In the first-best economy there are neither agency ($\Lambda(\cdot) = 0$) nor financial ($(\Phi, \phi) = (0, 0)$) frictions, management always chooses to exert effort, and the firm holds no cash and pays free cash flow out immediately. Because the economic environment is *iid* and the model homogeneous, there is a constant investment rate that maximizes firm value:[10]

$$q^{FB} = \max_i \frac{(1 - \tau_Y)\mu - g(i)}{r + \delta - i}. \tag{2.12}$$

[10] We assume $\mu < g(r + \delta)$ and $q^{FB} > l_K$ to have a well-defined problem.

In this economy, the classic Hayashi (1982) result equates average Q and marginal q to investment's marginal cost to solve for optimal investment:

$$g'(i^{FB}) = q^{FB} = \frac{(1-\tau_Y)\mu - g(i^{FB})}{r+\delta-i^{FB}}. \tag{2.13}$$

Finally, because our managers are relatively impatient, it is best to pay them immediately, leaving $P^{FB}(K,W) = q^{FB}K - wK$ to investors.

2.3.2 Cash Management With No Agency Conflict

In the absence of an agency friction ($\Lambda(\cdot) = 0$), management will always choose to exert effort. The only state variable is scaled cash holdings, $c = C/K$, implying firm value per unit of capital, $p(c)$, only depends on c, and the setup is similar to Bolton, Chen, and Wang (2011a). We follow their exposition when it is always optimal to refinance the firm.

Because of the fixed issuance cost, the firm will want to minimize instances of refinancing and will therefore only do so when resources reach zero. When refinancing, the firm receives a total issue amount of $f > 0$ per unit of capital. Because firm value is continuous before and after issuance, value matching at the refinancing boundary $c = 0$ holds:

$$p(0) = p(f) - \Phi - (1+\phi)f. \tag{2.14}$$

The right side is the firm's post-financing value less fixed issuance costs Φ and proportional financing costs ϕ. Because f is optimally chosen, smooth pasting equates the marginal value of the last dollar raised $p'(f)$ to one plus the marginal financing cost,

$$p'(f) = 1 + \phi. \tag{2.15}$$

Conversely, because holding cash in the firm is penalized at rate τ_C, the firm will distribute it to investors when abundant. Formally, let $\overline{c}$ denote this endogenous payout boundary and for $c > \overline{c}$ we have the equation $p(c) = p(\overline{c}) + (c - \overline{c})$. Because this equation holds continuously, the limit $c \to \overline{c}$ is summarized by the derivative

$$p'(\overline{c}) = 1. \tag{2.16}$$

Intuitively, at $\bar{c}$ the firm is indifferent between distributing and retaining one dollar, so the marginal value of cash must equal one. Since the payout boundary is optimally chosen, we also have the super contact condition holding at this point

$$p''(\bar{c}) = 0. \tag{2.17}$$

To summarize, incremental payouts dD occur when $c \geq \bar{c}$ and incremental financing dF is received when $c = 0$. Within these boundaries both dD and dF are zero and the dynamics of (2.1) and (2.4) imply, by Ito's lemma, that the evolution of scaled cash holdings is

$$dc_t = [(1-\tau_Y)\mu - g(i_t) + \tau_Y\delta + (r(1-\tau_C) - (i_t - \delta))c_t]dt + \sigma(1-\tau_Y)dZ_t. \tag{2.18}$$

The solution to investors' problem in (2.11) is in turn determined by the ordinary differential equation

$$\begin{aligned} rp(c) = \max_i\, & p(c)(i-\delta) + p'(c)[(1-\tau_Y)\mu - g(i) + \tau_Y\delta + (r(1-\tau_C) - (i-\delta))c] \\ & + \frac{1}{2}p''(c)\sigma^2(1-\tau_Y)^2 \text{ for } c \in [0, \bar{c}] \end{aligned} \tag{2.19}$$

subject to the boundaries (2.14), (2.15), (2.16), and (2.17) that jointly pin down the location and optimality of the refinancing and payout decisions concerning f and $\bar{c}$. The optimal rate of investment takes the form

$$g'(i) = \frac{p(c)}{p'(c)} - c \tag{2.20}$$

where its marginal cost is equated to its marginal benefit; namely, average Q, $p(c)$, adjusted for the marginal value of cash $p'(c)$, less the reduction in cash holdings c. If cash is scarce, its marginal value is high and then, for a given Q, investment is diminished. Cash holdings thus influence the choice of investment.

2.3.3 Agency Problem with Costless Refinancing

In the presence of an agency conflict yet the absence of external financing costs ($(\Phi, \phi) = (0, 0)$), episodes of refinancing are left indeterminate. Because cash held in the firm would incur a penalty, free cash flow is immediately paid to investors. The only state variable becomes managers' stake, $w = W/K$, making firm value per unit of capital to investors equal to $p(w)$. The setup is then similar to DeMarzo, Fishman, He, and Wang (2012) and following them the firm is liquidated when the contract is terminated.

Management will be terminated once their continuation utility hits zero (their outside option) because otherwise they would immediately consume private benefits. Hence, investors' liquidation payoff at this termination boundary is

$$p(0) = l_K. \tag{2.21}$$

Next, because investors can always compensate management with cash, it will cost at most one dollar to increase w by one dollar, implying $p_w(w) \geq -1$. But because termination is costly ex post it will be optimal to grow w at low values as quickly as possible by setting incremental payments dU/K in (2.8) to zero. As management is more impatient however ($\gamma > r$), at some point they will need to receive current payments. Formally, this payment boundary is the threshold where investors are indifferent between reducing their value by one dollar to pay agents one dollar immediately

$$p'(\overline{w}) = -1, \tag{2.22}$$

and because it is determined optimally it satisfies the super contact condition

$$p''(\overline{w}) = 0. \tag{2.23}$$

To summarize, $dU_t/K_t = 0$ within the payment and termination boundaries and $\beta_t = \lambda_K/((1-\tau_Y)\mu)$ for all t. The evolution of w that is derived from the optimal contract then follows from (2.1) and (2.8):

$$dw_t = (\gamma - (i_t - \delta))w_t dt + \frac{\sigma}{\mu}\lambda_K dZ_t. \tag{2.24}$$

With the dynamics of management's payoff given, the solution to investors' problem in (2.11) is concave (see DeMarzo, Fishman, He, and Wang (2012) for details) and takes the form

$$\begin{aligned} rp(w) = \max_i (1-\tau_Y)\mu - g(i) + \tau_Y\delta + p(w)(i-\delta) + p'(w)(\gamma - (i-\delta))w \\ + \frac{1}{2}p''(w)\left(\frac{\sigma}{\mu}\lambda_K\right)^2 \quad \text{for } w \in [0, \overline{w}] \end{aligned} \tag{2.25}$$

that is pinned down by the boundaries (2.21), (2.22), and (2.23) and where investment is determined by

$$g'(i) = p(w) - p'(w)w. \tag{2.26}$$

When choosing investment, investors internalize its effect on managers' incentives. For a given $w = W/K$, an increase in capital reduces management's effective claim on the firm and induces a more severe agency friction.

2.3.4 Complete Model

We now discuss the complete solution in (2.11). It nests the previous two models as special cases and for brevity we streamline its presentation and draw attention only to novelties. Common to the nested models, within the boundaries of the solution the firm finances itself internally and only affects management's incentives through their continuation payoff. Total firm value per unit of capital now depends on both scaled cash holdings and managers' stake, $p(c, w) + w$.

Boundaries

As before, the optimal contract specifies termination when managers' stake equals their outside option

$$p(c, 0) = l_K + l_C \times c, \text{ for all } c. \tag{2.27}$$

Thus regardless of the level of cash the firm will be liquidated when $w = 0$ because at this point managers will shirk ($e_t = 0$) and investors will terminate the contract.

At stakes above $w = 0$, firm value satisfies $p_w(c, w) \geq -1$ and, again, at the payment boundary investors will be indifferent between promising and paying managers one dollar

$$p_w(c, \overline{w}(c)) = -1, \text{ for each } c, \tag{2.28}$$

while the super contact condition determines the level of the boundary itself

$$p_{ww}(c, \overline{w}(c)) = 0, \text{ for each } c. \tag{2.29}$$

We emphasize that the payment boundary $\overline{w}(c)$ is now a function: investors will choose to raise or lower the threshold depending on the level of cash. For example, if cash holdings increase from a small amount, the value of the firm will rise as the likelihood of costly refinancing falls. Because firm value is now higher, it will be optimal to reduce the probability of inefficient termination. Thus when starting at a low cash level, investors will find it efficient to further raise $\overline{w}(c)$ along with c and shrink the likelihood of termination.

Next, we turn to the boundaries that determine the decisions of refinancing and payouts to investors. As before the firm refinances only when it runs out of cash, but the magnitude, $f(w) > 0$, is now state-dependent.[11] To see this, since firm value is continuous before and after equity issuance it implies that

$$p(0, w) = p(f(w), w) - \Phi - (1 + \phi)f(w), \text{ for each } w, \tag{2.30}$$

where smooth pasting satisfies

$$p_c(f(w), w) = 1 + \phi, \text{ for each } w. \tag{2.31}$$

The size of refinancing now depends on management's stake, w. It is natural to believe that investors would likely provide more funds to a firm with a shrinking agency conflict. This intuition implies that we should expect to see regions where $f'(w) > 0$.

Because holding cash is costly, the firm will pay out once holdings are sufficiently large. Since firm value must be equal before and after a payout, the exact amount obeys the equation $p(c, w) = p(\overline{c}(w), w) + (c - \overline{c}(w))$ for $c > \overline{c}(w)$, where we can see the payout boundary $\overline{c}(w)$ is now a function of w. And because it is continuous, the equation holds in the limiting case as $c \to \overline{c}(w)$ and implies the boundary

$$p_c(\overline{c}(w), w) = 1, \text{ for each } w, \tag{2.32}$$

and again optimality requires

$$p_{cc}(\overline{c}(w), w) = 0, \text{ for each } w. \tag{2.33}$$

Altogether, the payout boundary becomes dependent on management's stake, and the relation is not necessarily monotone for the following reasoning. It is possible in equilibrium that if w is sufficiently high it could be efficient to lower the cash threshold

[11] We can show refinancing is always optimal at zero cash holdings. When holdings become zero, the liquidation value of the firm is $l_K - w$. Because $p_w(0, w) \geq -1$ over $w \in [0, \overline{w}(0)]$ and with equality when $w = \overline{w}(0)$, we have $p(0, w) = p(0, 0) + \int_0^w p_w(0, w')dw' \geq p(0, 0) + \int_0^w [-1]dw' = p(0, 0) - w = l_K - w$.

for payouts, conceivably making a costly refinancing event more likely. Of course, at high w investors know managers incentives are well aligned and will be motivated to keep refinancing a distant event. This effect arises because our economy is dynamically consistent and is also novel to the complete model.

And finally, we require along the boundary curves $\overline{c}(w)$ and $\overline{w}(c)$ that super contact holds with respect to both states. Economically, this means that along a segment of $\overline{c}(w)$ it is optimal to not give managers current payments and similarly that along a part of $\overline{w}(c)$ that investors will not earn payouts. Of course, there could be segments on which $\overline{w}(c)$ and $\overline{c}(w)$ overlap where both managers receive payments and investors payouts and form a joint upper boundary. These last few technical conditions follow below.

To ensure that $\overline{c}(w)$ achieves super contact we first differentiate $p(c,w) = p(\overline{c}(w), w) + (c - \overline{c}(w))$ with respect to w which gives

$$p_w(c,w) = \frac{\partial p(\overline{c}(w), w)}{\partial w} - \frac{\partial \overline{c}(w)}{\partial w}, \text{ for each } c \geq \overline{c}(w). \tag{2.34}$$

Since the equation's right side is not a function of cash, taking a derivative with respect to c and letting $c \to \overline{c}(w)$ implies

$$p_{wc}(\overline{c}(w), w) = 0, \text{ for each } w. \tag{2.35}$$

Next, a similar idea starting from $p(c,w) = p(c, \overline{w}(c)) - (w - \overline{w}(c))$ but differentiating with respect to c and then w and letting $w \to \overline{w}(c)$ gives

$$p_{cw}(c, \overline{w}(c)) = 0, \text{ for each } c. \tag{2.36}$$

And finally from equations (2.35) and (2.36) it is evident that

$$p_{cw}(\overline{c}(w), \overline{w}(c)) = 0, \text{ for every } c \text{ and } w. \tag{2.37}$$

Summary

To summarize, events of refinancing dF, payouts dD, and payments to managers dU are zero within the boundaries and termination occurs when $w = 0$ regardless of cash holdings. Here, system dynamics are governed by

$$dc_t = ((1-\tau_Y)\mu - g(i_t) + \tau_Y\delta + [r(1-\tau_C) - (i_t - \delta)]c_t)dt + \sigma(1-\tau_Y)dZ_t \tag{2.38}$$

$$\text{and } dw_t = (\gamma - (i_t - \delta))w_t dt + \frac{\sigma}{\mu}\lambda(c_t)dZ_t. \tag{2.39}$$

Both resources and incentives vary with productivity and the optimal contract makes these two variables perfectly correlated. Because the drifts will differ in general, however, the stationary distribution will display the rich tradeoffs of the economic environment.

One such tradeoff arises from the interdependence of (2.38) and (2.39). The optimal contract sets $\mu e_t = \mu$ implying that cash will grow quickly. But as cash holdings accumulate, so does the severity of the agency friction, $\lambda(c)$. This raises the sensitivity of management's compensation to the underlying productivity shocks. So while growth in productivity benefits cash holdings and potentially delays a refinancing event, the likelihood of contract termination increases for a given w, making the firm riskier. Thus riskier firms will generally hold more cash, consistent with the empirical evidence in Acharya, Davydenko, and Strebulaev (2012).

In what follows we assume that $\beta p_{ww}/2 + p_{cw}$ and p_{ww} are nonpositive, conditions which we discuss further and verify numerically in Appendix B. The solution to (2.11), then, can be represented by the partial differential equation

$$\begin{aligned} rp(c,w) = \max_i\, & p(c,w)(i-\delta) + p_c(c,w)\left((1-\tau_Y)\mu - g(i) + \delta\tau_Y + [r(1-\tau_C) - (i-\delta)]c\right) \\ & + p_w(c,w)\left((\gamma - (i-\delta))w\right) + \frac{1}{2}p_{cc}(c,w)(\sigma(1-\tau_Y))^2 \\ & + \frac{1}{2}p_{ww}(c,w)\left(\frac{\sigma}{\mu}\lambda(c)\right)^2 + p_{cw}(c,w)\frac{\sigma^2(1-\tau_Y)}{\mu}\lambda(c). \end{aligned} \tag{2.40}$$

subject to incentive boundaries in (2.27), (2.28), and (2.29) that determine termination and the payment threshold to management, $\overline{w}(c)$; the resource boundaries of (2.30), (2.31), (2.32), and (2.33) that locate the position and ensure the optimality of the refinancing and payout decisions, $f(w)$ and $\overline{c}(w)$; as well as the mixed boundaries given by (2.35), (2.36), and (2.37). We detail our computational method to solve this problem in Appendix A. A key novelty relative to Achdou, Han, Lasry, Lions, and Moll (forthcoming) is that the shape of state space is solved jointly with (2.40).

The optimal investment decision is determined by

$$g'(i) = \frac{p(c,w) - p_w(c,w)w}{p_c(c,w)} - c \tag{2.41}$$

and now reflects several margins: the gain in value, $p(c,w)$, less the detrimental change to managers' incentives, $p_w(c,w)w$, adjusted for the marginal value of cash, $p_c(c,w)$, and its reduction, c. More generally, the decision deepens the link between cash holdings, compensation, and investment. Empirically, increasing long-term incentive plans (LTIP) raises investment (Larcker (1983) and Glover and Levine (2017)) as do cash holdings, a result more or less established in Fazzari, Hubbard, and Petersen (1988) and subsequently refined by a large literature.

Distortions and The Joint Upper Boundary

A key tradeoff balances greater cash holdings c with a greater agency friction $\lambda(c)$ while considering both parties' interests. Ideally, the firm would locate where both financing and agency frictions and their value distortions are minimized, which respectively correspond to where $p_c(\overline{c}(w), w) = 1$ and $p_w(c, \overline{w}(c)) = -1$.

A remarkable outcome is that in spite of the problem's complexity, our setup produces a simple, intuitive tradeoff along the joint optimal boundary, a fact which we summarize in the following proposition:

Proposition (Tradeoff Along the Joint Upper Boundary)**.** *Consider a marginal change along the joint upper boundary from* $(\overline{c}(w), \overline{w}(c))$ *to* $(\overline{c}(w) + dc, \overline{w}(c) + dw)$*, then the rate of change across this boundary is equal to*

$$\frac{dw}{dc} = -\frac{r\tau_C}{\gamma - r} < 0 \tag{2.42}$$

Proof. See Appendix B. ■

At the boundary at which investors receive payouts and managers payments, the slope equals the ratio of marginal costs of retaining cash to withholding payments to managers. Even though the value distortions have dissipated ($p_c(\cdot) = 1$ and $p_w(\cdot) = -1$), the lingering sources of inefficiency, the cash tax penalty and managers' impatience, survive and determine the optimal tradeoff faced by firms.

More specifically, if investors decide to hold an additional amount dc of cash, they will bring managers' payment boundary inwards by $\frac{r\tau_C}{\gamma - r}dc$. Accordingly, the proposition suggests that the rates and magnitudes at which managers or investors are differentially paid are informative about the relative size of these underlying frictions.

More generally, movements away from the joint upper boundary either increase cash's marginal value, $p_c(\cdot) > 1$ as costly refinancing becomes more likely, or raise the marginal value of an additional dollar promised to managers, $p_w(\cdot) > -1$, as it is becomes more valuable to avoid costly termination. Therefore as either c or w decrease, the marginal costs of not paying out shareholders or not currently paying managers fall and change the tradeoff faced along the boundary.

Mathematically this means the curvature of the boundary on either side of the joint upper boundary reflects marginal changes in relative distortions. Shifts away from $(\overline{c}(w), \overline{w}(c))$ and along $(\overline{c}(w), w)$, for example, are informative about the distortion attributed to the agency friction. The model can novelly be used to measure these distortions.

In our specified setup, the tradeoff is exactly linear at the joint upper boundary and a different specification, like decreasing returns to scale, could relax this linearity. The general insight that the curvature of the boundary is informative about the underlying magnitudes of distortions, however, would remain.

Aggregation

With the description of firm behavior complete, we now describe the stationary distribution of firms. While the payout, payment, and refinancing boundaries are not absorbing, the termination boundary is. Because of this, every firm will eventually fail and in order to study a stationary distribution we therefore allow entry. The exit rate, moreover, is an salient equilibrium object that governs the severity of the agency friction.

Each firm is described by its current state (c, w), and therefore the density of firms is defined over this state space. The non-stationary distribution at time t, $h(c, w, t)$, satisfies the Kolmogorov forward equation

$$\frac{\partial h(c,w,t)}{\partial t} = \varphi(c,w)m + \mathcal{A}^* h(c,w,t), \tag{2.43}$$

where $\mathcal{A}^* h(c,w,t)$ is the adjoint of the infinitesimal generator of the bivariate diffusion process (dc_t, dw_t).[12] By construction, this generator contains the rates of exit that occur along the termination boundary $w = 0$. To ensure a stationary mass of firms, we add a product of an entry rate m and an entry mass $\varphi(c,w)$ that integrates to one.

We pin down the entry rate in the stationary distribution with the normalization that the total mass of firms is a constant equal to one: $\int_0^{\overline{c}(w)} \int_0^{\overline{w}(c)} h(c,w)dwdc = 1$. After this normalization notice the left side of (2.43), twice integrated, is zero and we can then rearrange it for the stationary entry rate, which by construction equals the exiting mass of firms,

$$m = -\int_0^{\overline{c}(w)} \int_0^{\overline{w}(c)} \mathcal{A}^* h(c,w)dwdc. \tag{2.44}$$

When a firm's contract is terminated, a new, replacing firm's cash holdings is drawn from a distribution with positive support. The entrant, however, also starts with a new continuation payoff, w_0, that is determined by the bargaining power of agents and investors. We specify initial conditions during our calibration in the next section.

2.4 Model Calibration and Analysis

Having characterized the solution to the model, we now calibrate it before turning to study the solution's properties. After the calibration and as before, we build to the complete model by revisiting the predictions of its nested models. We then define the measurement of financial and agency distortions.

2.4.1 Calibration

Our calibration is summarized in Table 2.1. It is split into externally- and internally-calibrated parameters that target informative data moments. Our empirical environment contains only US public firms, as agency frictions are likely to be present among

[12] Specifically, $\mathcal{A}^* h(c,w,t)dt = -\mathbb{E}_t[dc]h_c(c,w,t) - \mathbb{E}_t[dw]h_w(c,w,t) + \frac{1}{2}\mathbb{E}_t[(dc)^2]h_{cc}(c,w,t) + \frac{1}{2}\mathbb{E}_t[(dw)^2]h_{ww}(c,w,t) + \mathbb{E}_t[(dcdw)]h_{cw}(c,w,t)$.

them, and report the details of our widely-used Compustat and Execucomp data samples in Appendix F.

We begin by calibrating our external parameters by setting the tax rate on corporate income to $\tau_Y = 30$ percent, the interest rate to $r = 4$ percent, and the depreciation rate of assets to $\delta = 8$ percent, all common values in the literature. We choose to model a tax penalty for holding cash in the firm. Two relevant sections of the IRS tax code are Section 531 on the accumulated earnings tax and Section 541 on undistributed personal holding company income. Both sections impose the same penalty rate and we accordingly use $\tau_C = 20$ percent.

We specify a smooth adjustment cost technology as we are interested in the model's long-run properties that takes the quadratic form

$$g(i) = i + \frac{\theta}{2}(i - \delta - z)^2, \tag{2.45}$$

where θ measures the magnitude of the adjustment cost and z is an exogenous expansion rate that locates the function. Following Hall (2001), we interpret the parameter θ as a doubling time of capital. He uses either 2 or 8 years for his upward adjustment cost and 20 or 80 years for his downward adjustment cost. Because financial and agency frictions will lower investment rates below first-best and potentially the depreciation rate of capital, we assume an exogenous expansion rate z to match the average investment rate in the data and generate long-run capital growth. Our choices are $\theta = 6$ and $z = 9$ percent.

Next, we turn to the costs of refinance. Beginning with the seminal work of Jensen and Meckling (1976) and Myers and Majluf (1984), subsequent literature has tried to estimate indirect costs, like asymmetric information and incentive costs, and direct costs, underwriting fees and dilution for example. Estimates vary across studies: Calomiris and Tsoutsoura (2013) argue a 3 percent decline in the price of equity in response to a seasoned equity offering is reasonable but can be as high as 15 percent for smaller firms in totality when accounting for all costs; and Altinkiliç and Hansen (2000) show that the majority of costs for a seasoned offering are variable, ranging from 4 to 6 percent depending on issue size, with fixed costs slightly below half a percent. Informed by them, we impose $\Phi = 0.5$ percent and $\phi = 5$ percent.

In the event of contract termination, we assume the firm is liquidated. In a recent study of recovery rates within bankruptcies, Kermani and Ma (2020) estimate values between 33 and 46 percent for all assets. Chen (2010) structurally estimates average recovery rates for bondholders to near 40 percent. We assume a $l_K = 40$ percent recovery rate for capital and the perfect recovery of cash, $l_C = 1$.

Internal Calibration: Averages

We internally calibrate our remaining six parameters $(\mu, \sigma, \gamma, \lambda_K, \lambda_C, \psi)$ to features of the stationary distribution that clearly map model to data. This distribution encodes all information about optimal policy functions and is therefore an ideal target for calibration. Specifically, we target moments of free cash flow, compensation, and cash holdings as well as the shape of its state space that determine the frequencies and magnitudes of payouts, refinancing, and termination.

The incremental return on capital, μ, directly influences the mean rate of free cash flows. We set $\mu = 0.17$ to match this average.

Specialists' time rate of preference is $\gamma > r$. Its value influences the length of the interval $[0, \overline{w}(c)]$, as greater impatience (higher γ) requires sooner current payments and lowers $\overline{w}(c)$. In reality managerial compensation, while easily measurable, is complex as it contains salary, variable bonuses, long-term incentive plan contributions, and stock and options, the timing of which can also follow a complicated structure. As a resolution, we convert the stock of expected discounted future compensation, W, to a flow by multiplying it by γ, avoiding the subjective calculation necessary to evaluate (2.7) and effectively using the flow of compensation in the data to match the optimal contract's promising-keeping condition and track management's stake in the firm. Altogether, we set the parameter by calibrating to average compensation in the data and correspondingly choose $\gamma = 0.048$.

Finally, we set the agency costs across both capital and resources as $\lambda_K = 0.04$, which influences the variation in management's continuation utility and the probability of contract termination, and $\lambda_C = 0.09$, which influences cash's marginal value and therefore average cash holdings.

Recall that failure is source of inefficiency in the model and provides discipline to managers who under individual rationality would prefer to remain in the firm. The average termination rate in the model is m from (2.44). We target this rate with $\sigma = 35$

percent to match the two percent average default rates of public firms over 1993 to 2017 (Boualam, Gomes, and Ward (2020)). Thus our volatility parameter is targeted at the frequency of events which are determined by the boundaries rather than, say, the cross-sectional dispersion in investment rates.

We factor the entry mass into a conditional and marginal distribution, $\varphi(c_0, w_0) = \varphi_c(w_0|c_0)\varphi_c(c_0)$ and assume initial scaled cash holdings, c_0, draws from a log-normal distribution with mean $0.15 - \sigma^2/2$ and standard deviation σ, which generates an entrant's cash holdings close to the model's average refinancing size. Given the cash draw, c_0, the distribution of w_0 is degenerate and the value of initial w_0 comes from an assumption of managers' relative bargaining power. From wage responses to news, Taylor (2013) structurally estimates relative bargaining power to be equally split between shareholders and the chief executive and, accordingly, we pick $\psi = 50$ percent.

Internal Calibration: Frequencies and Magnitudes

Next, we describe how we construct the model's frequencies and magnitudes of refinancing and payouts to closely match their empirical construction. In the data, we form an indicator for a firm for whether it had ever, over the course of an entire year, had refinanced or paid out and simply average over these indicators to estimate frequencies. We do not record multiple events of the same firm within a year.

Our motivation for calculating the payout frequency in the model, then, follows from the question: Given the stationary mass at a point, (c, w), what fraction of firms would be expected to breach the payout boundary following a productivity shock, Δz, given at an annual rate? Given the shock, cash holdings and managers' stake move by $\Delta c = \mu_c(c, w) + \sigma_c \Delta z$ and $\Delta w = \mu_w(c, w) + \sigma_w(c)\Delta z$, respectively, where $\mu_c(c, w)$, σ_c, $\mu_w(c, w)$, and $\sigma_w(c)$ are the annualized drifts and volatilities of (2.38) and (2.39). Given these moves, we calculate

$$\text{Refinancing Rate: } \mathbb{E}[\mathbf{1}\{\Delta c + c < 0\}|c, w] = \mathcal{N}\left(-\frac{\mu_c(c, w) + c}{\sigma_c}\right), \tag{2.46}$$

$$\begin{aligned} \text{Payout Rate: } & \mathbb{E}[\mathbf{1}\{\bar{c}(w + \Delta w) < \Delta c + c\}|c, w] \\ & = \mathbb{E}\left[\mathbf{1}\{\Delta z > (\bar{c}(w + \mu_w(c, w) + \sigma_w(c)\Delta z) - c - \mu_c(c, w))/\sigma_c\}\right], \end{aligned} \tag{2.47}$$

where $\mathcal{N}(\cdot)$ is the cumulative standard normal distribution and we compute the rate of payouts with Gaussian quadrature while accounting for the correlation between c and w. Refinancing size is simply $\mathbb{E}[f(w)|w]$ and is consistent with our empirical construction. These objects are conditional on (c, w) and we integrate over them with the stationary density to calculate refinancing and payout statistics.

Internal Calibration: Summary

The summary of the internal calibration is tabulated in Table 2.2. The model matches well the data's average levels of cash, compensation, investment, free cash flow, and entry/exit and refinancing rates.

One feature of the data that the model has difficulty in matching is refinancing size. When it occurs in the data, it raises a much larger amount on average than in the model. Small firms are known to raise a lot more upon refinancing (Fama and French (2005)) and so allowing for decreasing returns to scale would help the model in this dimension.

The payout rate of the model, moreover, is lower than in the data. Payouts in the data are measured using common dividends and repurchases. In the model, payouts are more akin to special one-time dividends and repurchases, and so its frequency will naturally be lower since dividend policies are known to be quite persistent (Lintner (1956)). Additionally, the decision to return cash to investors would depend on the rate of return to the firm's investment. Here, decreasing returns to scale might also help as they would cause larger firms to have relatively lower investment returns and make payouts more appealing.

2.4.2 Solution to Cash Management with No Agency Conflict

Figure 2.1 plots the solution to the optimal cash management policy without an agency conflict being present. In Panel A, enterprise value is plotted in blue, $p(c) - c$, over the domain of scaled cash holdings, $c = C/K \in [0, \overline{c}]$, where $\overline{c}$ is the payout boundary. The red dashed line depicts marginal financing costs, $\Phi + \phi c$, and its tangency with enterprise value determines the refinancing size f.

Because refinancing is always preferred to liquidation, firm value is above l_K when cash is zero. It is concave in this region, reflecting the precautionary motive induced by the prospect of costly refinancing. As c grows and eventually reaches $\overline{c}$, the slope of enterprise value becomes zero. Concurrently, cash's marginal value is initially above one but converges to one.

Panel B plots the investment rate, $i = I/K$, as a function of cash holdings. The decision is influenced by the marginal value of cash that reflects financial frictions. As this marginal value grows the firm reduces investment to tilt the drift of cash's evolution upwards and reduce the likelihood of costly refinancing.

The stationary density of cash holdings is depicted in gray. Intuitively, the mass of firms above the refinancing size f dominates the mass below and cluster near the upper boundary. Refinancing is costly and firms avoid this by accumulating cash and making investment decisions to promote this goal.

2.4.3 Solution to Agency Problem with Costless Refinancing

Figure 2.2 depicts the solution to the agency model in the absence of external financing costs. Panel A shows the value function of investors, $p(w)$, under a contract of commitment and over the domain of managers' stake, $w = W/K \in [0, \overline{w}]$, where $\overline{w}$ is the payment boundary.

Because termination is ex post inefficient, investors become averse to fluctuations in w. Investor's value function, $p(w)$, is thus concave as investors internalize the risk that the optimal contract places on management's actions. At $w = 0$ the contract is terminated and investors receive the liquidation value l_K. Upon termination, a new firm is drawn depending on managers' bargaining power, ψ.

Two forces drive the shape of $p(w)$. Initially as w grows from zero, the agency friction falls and agrees with investors' desire to weaken it to avoid liquidation, $p(w)$ thus increases. To further alleviate the agency friction, investors must promise a larger and larger share of the firm to management, and therefore $p(w)$ eventually declines reflecting this wealth transfer. As w increases, the slope of the value function declines and eventually becomes -1 at the payment boundary, the point at which investors are indifferent between promising and paying managers one dollar.

Panel B plots the investment rate. To understand this graph it is best to think of the firm's total value being determined in part by investors' capital, K, and in part by management's ability to run the firm and their stake in the firm, W. Investment is costly to investors as it lowers management's effective ownership of the firm, $w = W/K$, and hence induces a greater agency friction. The cost of a greater friction is much larger at low levels of w and this is why investment is reduced. As the friction wanes with growing w, investment waxes along with it.

Finally, the stationary density of management's stake is shown in gray. Because liquidation is costly, the optimal contract sets $du = 0$ to grow w as quickly as possible. Because of management's impatience, however, du is eventually set to be positive at the payment boundary, $\overline{w}$. The density accumulates at this upper bound.

2.4.4 Solution to Complete Model and its Properties

The complete model retains the properties of the two nested models previously described. Investment, for example, is generally reduced as w or c fall from the payment or payout boundaries. The optimality of its solution similarly requires investors' value function to possess certain properties prescribed by the equilibrium contract as well as the decisions that determine the boundary conditions. In particular, the derivatives of the value function to each state, w and c, should be monotone decreasing functions, respectively, and the second derivatives should equal zero at the upper boundaries.

We depict the accuracy of these properties in Figure 2.3. Panels A and B report the first own-derivatives of a state across three percentiles of the other state's marginal density. Both the marginal cost of compensation $p_w(c, w)$ and the marginal value of cash $p_c(c, w)$ are decreasing functions, implying own-state concavity of the value function. As they approach their respective boundaries, the rate of change of these derivatives fall and approach zero. A notable difference between the complete model and the agency model with costless refinancing is that $p_w(c, w) < 0$ for all w here and therefore the contract is renegotiation-proof.

Next, in Panels C and D we plot the super contact condition associated with the payout and payment boundaries. If the decision is optimal, they should both be uniformly zero across the entire boundary. In general they are very close, although in the tails of the distribution of the state in question the magnitude of the deviation from zero grows. These deviations visually overstate the impact on the model's predictions, as they are concentrated over states on which the equilibrium stationary distribution puts little mass, as depicted by the marginal densities in gray. Sensitivity analysis confirms that the quantitative predictions of the model are robust to local changes in boundary curves.

Figure 2.4 displays the complete solution. Panel A shows the value function that is solved jointly with the state space. The boundary touches both zero axes for cash holdings and managers' stake. Remaining on a curve while moving in to the interior, the boundary curve eventually reaches the joint upper boundary, marked by a black square. The rate of change of either boundary, $\overline{c}(w)$ or $\overline{w}(c)$, reflects the current marginal values of states, $p_c(c, w)$ and $p_w(c, w)$, that in turn reflect the underlying financial and agency frictions.

Intuitively, the rise of c from zero coincides with a higher firm value and a greater likelihood of termination, making it efficient to raise the payment boundary $\overline{w}(c)$. At some point, however, the cost of termination falls as cash holdings and liquidation value have grown, bringing the payment boundary back in. In addition, as w moves away from zero, firm value increases and reduces the probability of termination. The expected return on investing capital within the firm is thus higher and improves the motive to push the payout boundary out farther. When managers' incentives have improved enough, both party's interests are well-aligned, and it becomes optimal to lower the payout boundary, reflecting our dynamically consistent contract whereby managers share investors interests. In general, the overall shape is determined by these competing effects.

The refinancing curve, $f(w)$, is plotted as a dotted black line. This function captures the disciplinary role of markets on management. In equilibrium, management understands that refinancing will be dependent on the firm's history up to that point and therefore this will change their behavior ex ante. As expected, refinancing does increase with w. As w rises from zero, $f(w)$ initially grows rapidly, but then plateaus and eventually falls.

Why it falls is reminiscent of Zwiebel (1996), who argues that managers voluntarily set debt to restrict themselves. Analogously, in our model management internalizes refinancing's effect on their incentives. They realize that with already good incentives, a larger issue causes unnecessarily greater risk to investors as $\lambda(c)$ and the risk of termination grow with c for a given w, and refinancing $f(w)$ thus falls. In general, the quantitative effects of increasing $f(w)$ and managers voluntarily restraining themselves determine the appearance of the curve.

The stationary density of firms is graphed in Panel B. The density approximately tracks out a path from zero cash holdings to the payout boundary, a track far above the termination boundary. Its particular gradient arises from the equilibrium investment decisions that drive the state equations. Because the model is stochastic the mass of firms are spread out along this track, causing some firms to fail.

2.4.5 Measurement of Distortions

The traditional approach to measuring distortions typically compares market values and investment across first-best and second-best outcomes. Market values, however, rely on an accurate discounting of future cash flows that are very influenced by hard-to-measure discount rates. Fully understanding investment, moreover, requires a good proxy for marginal q, an object notoriously hard to estimate (Erickson and Whited (2000)). First-best, furthermore, may not be a reasonable benchmark as it is unlikely to be attained in practice as financial and agency frictions do exist.

The Proposition of Section 2.3 defines the slope of the joint upper boundary and implies the existence of a *second-best frontier* that is tangent to this boundary. We therefore use the proposition to measure distortions relative to this frontier and form measurements based on quantities rather than on prices.

We construct an orthogonalized definition of the *agency* distortion by evaluating the distance between the frontier and the *payout* boundary, weighted by the mass of firms across this distance. We use a similar definition for the financial distortion. We denote the frontier as a function of state by $\mathcal{F}(c)$ and $\mathcal{F}^{-1}(w)$ and in Appendix B we derive the formal definitions

$$\text{Agency Distortion: } \mathbb{E}[\Delta c\mathbf{1}\{\overline{c}(w) < c + \Delta c < \mathcal{F}^{-1}(w)\}|c, w], \tag{2.48}$$

$$\text{Financial Distortion: } \mathbb{E}[\Delta w\mathbf{1}\{\overline{w}(c) < w + \Delta w < \mathcal{F}(c)\}|c, w]. \tag{2.49}$$

Similar to the frequency calculations from before, they are based on the annualized drifts and volatilities of (2.38) and (2.39) and, as they are functions of (c, w), we integrate over them with the stationary density to calculate the economy's average agency and financial distortions.

Why do we measure the agency distortion by the distance from the payout boundary? Compare a firm at the joint upper boundary $(\overline{c}(w), \overline{w}(c))$ to one on $(\overline{c}(w), w)$ where $w < \overline{w}(c)$. For this firm, the financial friction is due only to the holding cash penalty, $r\tau_C$, since $p_c(\overline{c}(w), w) = 1$, so the only distortion from the second-best frontier can be attributed to agency. The idea is akin to holding using a supply shock to identify a demand elasticity.

Put differently, payouts which occur between $\overline{c}(w)$ and $\mathcal{F}^{-1}(w)$ are *required* by investors to compensate them for holding a firm with an acute agency friction. Analogously, current payments which occur between $\overline{w}(c)$ and $\mathcal{F}(c)$ are *promised* to managers to operate a firm with a scarce cash holdings. These average distortions thus provide an economic interpretation of how much investors are compensated on average to remain invested in an agency-laden firm or managers to remain operating a cash-poor one.

Figure 2.5 illustrates the approach to measuring the agency distortion. It depicts the top-down view of the stationary density. The second-best frontier, the dashed-dot line, bounds the state space from above and produces a tangency at the joint upper boundary marked by the black square. Given a firm at (c, w) and the dynamics of (2.38) alone (recall our distortions' definitions are orthogonalized), we can evaluate the probability that this firm will receive a shock and cross the payout boundary $\overline{c}(w)$. For a given w, we truncate the likelihood of shock realizations beyond $\mathcal{F}^{-1}(w)$ as they surpass second-best outcomes. Since the shock changes cash holdings (or compensation), we are measuring quantities.

To sum up, we provide novel, theoretically-consistent measurements of financial and agency distortions and show that payouts and compensation, two readily observable variables, can evaluate their magnitudes. Common to the literature in general, the measurements of these distortions are conditional on a model, as they are in Hsieh and Klenow (2009), and an alternative model would lead to different estimates. In Hsieh and Klenow (2009), they partially control for this shortcoming by comparing the same model across countries and evaluate relative distortions. A different yet also valid approach would be to pick a model and compute the ratio of payouts to payments. This ratio mimics an index of the *relative* severity of agency to financial distortions and, usefully, can be constructed for one industry in one country. Evaluating changes in this index are more robust and less affected by model misspecification.

2.5 Model Analysis and Evaluation of Market Discipline

With all objects defined and the model solved, we now focus on two forms of model analysis. First, we contrast impulse response functions across different calibrations. Recall that market discipline is an indirect force which we model by having only local shocks affect management's continuation utility in (2.8). Dynamic analysis allows us to visualize and quantify this force by showing that firms, even with the same initial (c, w), evolve differently over time.

Second, we use steady state analysis pioneered in Hopenhayn (1992) by observing how the stationary density of firms shifts in response to a change in a model parameter. We use it here to understand how changes in industrial structure, whether in the tax code or in the severity of a deep agency friction, affect the observable characteristics of firms and the distortions present in the economy. Though these stationary densities remain constant through time, they do so by the neutralizing effects of entry and exit, of firm growth and contraction. This analysis is therefore useful for understanding adaptation in ex ante behavior as it captures the long-run effects of these structural changes.

2.5.1 Evaluation of Market Discipline

We assess market discipline's quantitative effect in the following way. We first solve two economies that differ by a parameter. For illustration we contrast the fixed cost of refinance, setting Φ to 50 and 150 basis points. We then choose an initial pair of states that defines a firm, (c_0, w_0) and let investment, scaled cash holdings, and managers' stake evolve in the absence of future shocks and track the evolution of a firm's policy. Related, an impulse response function maps the dynamics of firm policies from the steady state after experiencing a normalized shock, typically. Because our steady state is composed of a distribution of firms, our notion here is the natural extension of an impulse response function to an arbitrary firm in the steady state distribution.

Our firm is initialized with scaled cash holdings near the average refinancing amount in data, $c_0 = 15$ percent of net assets, and a level of managerial compensation equal to the average in the data, γw_0, equal to 1.3 percent of net assets. The response functions are depicted in Figure 2.6. As expected, both scaled cash holdings and managers' stake start at the same level. Over time, however, they diverge because firm value and the state space differ.

Cash holdings are more quickly accumulated in the high Φ economy, ending up nearly 5 percentage points higher. Managers, however, end up with a smaller ownership stake in the firm. As cash is more valuable in the high Φ economy, the probability of refinancing within the year falls, but the likelihood of payouts is largely unaffected. Effectively managers, not investors, bear the change in fixed cost. That in turn predicts that distortions attributed to agency are now greater and basically only become quantitatively similar after 20 quarters, when nearing the joint upper boundary. Finally, financial distortions are roughly similar, as the effects of a greater fixed cost of refinance have been offset by a more rapid accumulation of cash.

Altogether, a small change in fixed cost of refinance can have a dramatic impact on managers' compensation, firms' cash holdings, and the agency distortion in the economy. This underscores that market discipline has a material force on economic dynamics.

2.5.2 Steady State Analysis

The results of this form of analysis are summarized in Table 2.3. For convenience, the first column restates the benchmark moments targeted in the internal calibration. It also includes estimates of average distortions and the ratio of their magnitude relative to the benchmark.

Our benchmark calibration has firms paying out 2.54 percent of firm assets per year out as a result of agency frictions. Managers are compensated by an additional 0.27 cents per dollar of assets to operate riskier firms with low cash ratios. We find agency frictions to be nearly ten times more severe than financial ones when measured with our model's quantities. The intuition is that financial frictions can simply be offset by accumulating cash (Bates, Kahle, and Stulz (2009) document an abnormally high accumulation rate since the 1990s). An accumulation of cash, however, is double-edged as it exacerbates the agency conflict, which requires payouts to investors (see Farre-Mensa, Michaely, and Schmalz (2014) for corroborating empirical evidence connecting payouts and agency).

The scenario in Column (2) raises average productivity, μ, to 18 percent. The firm now as a whole is more profitable and managers are correspondingly paid more and terminated less. The return to investment rises and the firm holds more cash, reducing the frequency of external finance. These patterns are consistent with a boom. In this economy, cash holdings and capital grow, implying that so do agency frictions. The model therefore predicts agency conflicts to be procyclical and financial frictions countercyclical.

In Column (3) we lower the corporate tax rate from 30 to 21 percent, matching a change enacted in 2017. The tax cut raises average asset productivity and strengthens the precautionary savings motive for holding cash. Although reducing corporate taxes is widely believed to lead to stimulating investment, we find that the effect is only modest as it raises the investment rate not even one percentage point. Investors instead prefer to allocate the additional free cash flow to managerial compensation to alleviate agency frictions and to distributing cash to themselves.

The change of the fixed refinancing cost, Φ, appears in column (4). Frequencies of payout and refinancing fall, similar to the boom-like patterns of column (2), but markets refinance on more stringent terms and as a result they terminate more firms. Greater cash holdings effectively offset a potentially aggravated financial friction. But managers must now operate more cautiously and with more resources that can potentially be squandered. Investors thus require more compensation for a larger agency conflict, as can be seen by an agency distortion growing from 2.54 to 2.92.

A Preliminary Proposal to Reduce Agency Frictions

Columns (3) and (4) collectively mimic a corporate income tax cut combined with the introduction of a tax of one percent on the instance of refinancing. We combine both changes in column (5). As before, a higher Φ reduces refinancing rates and increases the probability of termination, a change which alone would require a greater compensation for agency conflicts demanded by investors. Combining this change with lower corporate taxes, however, raises average cash flows that allows investors to reallocate a portion of them towards rewarding managers and alleviating agency conflicts.

In the final column (6) we initially lower the agency friction attributed to cash from 0.09 to 0.05 and then lower managers' bargaining power to match the entry/exit rate of column (6). Several rows across columns (5) and (6) look similar quantitatively. And relative to the benchmark case the frequency of refinancing has fallen and conditional on it happening, its average size has grown.

Altogether, an economy that implements the tax proposal above generates an economy that mimics one with a relatively less severe agency friction. This is an imputed result of market discipline. The relative values of average distortions are near identical.

In its current state, however, the analysis shown here is only suggestive and presents a tradeoff whereby the overall effects on agency and financial frictions must be weighed. Of course a fuller and potentially general equilibrium analysis that includes a government budget constraint would be required to be more confident in prescriptions for policy. But we find it interesting nonetheless and leave a more analysis to future work that we discuss next in the conclusion.

2.6 Conclusion

We quantitatively evaluate the fundamentally important question of the degree to which investor and managerial incentives are aligned and the role markets play in attaining firm value maximization. We formalize the notion of market discipline whereby markets, even though tapped intermittently, invisibly guide management's use of resources.

Our quantitative model clarifies the role of markets in affecting a wide range of firm policies, from cash holdings, investment, payouts, compensation, to whether to refinance a firm or let it fail. We also derive a novel, general formula that shows how investor payouts and managers' compensation are informative about the underlying distortions of costly external finance and agency conflicts. Our benchmark calibration has firms paying out 2.54 percent of firm assets per year out as a result of agency frictions and managers are compensated by an additional 27 cents per dollar of assets to operate riskier firms with low cash ratios, implying agency frictions are nearly 10 times more severe than financial ones.

While novel, our analysis necessarily omits some features that we believe are important. First is decreasing returns to capital. The model-data fit on statistics of refinancing and payouts would be expected to improve with this amendment. Yet another useful extension would be clearly formulate capital structure. Our current setup works best to describe large firms that are not overly indebted. Other improvements that could be equally important include time-variation in aggregate states. Lustig, Syverson, and

Van Nieuwerburgh (2011) partially attribute the rise in the disparity across executive compensation to changes in executive's outside options. Bolton, Chen, and Wang (2013) entertain a model where market conditions fluctuate and influence the costs of external financing over time. One last extension would be to model takeovers, board composition, and competition, as these likely influence managers' behavior. We leave these variations on our benchmark model to future work.

Table 2.1: **Variable Definitions and Calibration**

Note: This table defines model variables and the values of the benchmark calibration discussed in Section 2.4. All parameters are annualized.

Variable	Symbol	Parameter	Symbol	Value
Capital Stock	K	Agency Costs	(λ_K, λ_C)	(0.04,0.09)
Cash Holdings	C or $c = C/K$	Average Productivity	μ	0.17
Managers' Stake	W or $w = W/K$	Volatility of Productivity	σ	0.35
Investors' Value	$P(\cdot)$ or $p(\cdot) = P(\cdot)/K$	Depreciation	δ	0.08
Investment	I or $i = I/K$	Capital Adjustment Cost	θ	6
Adjustment Cost	$G(\cdot)$ or $g(\cdot) = i + G(\cdot)/K$	Exogenous Expansion Rate	z	0.09
Cumulative Productivity	A	Recovery Rates	(l_K, l_C)	(0.4,1)
Cumulative Free Cash Flow	Y	Interest Rate	r	0.040
Cumulative Payments	U	Management's Discount Rate	γ	0.048
Cumulative Payouts	D	Refinancing Costs	(Φ, ϕ)	(0.005,0.05)
Cumulative Refinancing	F	Corporate Tax Rate	τ_Y	0.30
Cumulative Refinancing Costs	X	Penalty on Cash Holdings	τ_C	0.20
Termination Time	τ	Managers' Relative Bargaining Power	ψ	0.5
Payment Boundary	$\overline{w}$ or $\overline{w}(c)$	Average Entrant Cash Holdings	$\mathbb{E}[c_0]$	0.15
Payout Boundary	$\bar{c}$ or $\bar{c}(w)$			
Refinancing Size	f or $f(w)$			

Table 2.2: **Internal Calibration Targets**

Note: This table reports averages and percentiles of several variables targeted in the data by the model's stationary distribution. The data annually cover the period from 1993 until 2017 and definitions are in Appendix F. Data variables are winsorized across all firm-years by 5 percent at the upper and lower tails except for refinancing size which is only winsorized at the upper tail. Percentiles are from the 25th, 50th, and 75th breakpoints. In the model, continuous variables are scaled cash holdings, $c = C/K$, compensation, $\gamma w = \gamma W/K$, investment $i = I/K$, and free cash flow $\mathbb{E}_t[dY]/K + \tau_Y \delta$. Indicator variables are payout rate (2.47), entry/exit rate (2.44), refinancing rate (2.46), and refinancing size which is $f(w)$.

	Model	Data			
	Mean	Mean	P25	P50	P75
Cash Holdings	21.8	22.7	2.1	7.8	26.2
Compensation	1.4	1.3	0.3	0.7	1.6
Investment	8.1	8.7	2.8	5.5	10.9
Free Cash Flow	4.0	4.8	-0.2	7.8	14.6
Payout Rate	37.9	53.2			
Entry/Exit Rate	1.2	2.0			
Refinancing Rate	16.4	16.8			
Refinancing Size	14.4	51.5	12.4	24.8	58.9

Table 2.3: **Steady State Analysis (Annual)**

Note: This table reports averages under the stationary density from various calibrations of the model. Variables are scaled cash holdings, $c = C/K$, compensation, $\gamma w = \gamma W/K$, investment $i = I/K$, and free cash flow $\mathbb{E}_t[dY]/K + \tau_Y \delta$. Indicator variables are payout rate (2.47), entry/exit rate (2.44), refinancing rate (2.46), and refinancing size which is $f(w)$. Agency and financial distortions are computed respectively in (2.48) and (2.49).

					$\tau_Y = 0.21$	$\lambda_C = 0.05$
	Benchmark	$\mu = 0.18$	$\tau_Y = 0.21$	$\Phi = 0.015$	$\Phi = 0.015$	$\psi = 0.03$
	(1)	(2)	(3)	(4)	(5)	(6)
Cash Holdings	21.8	25.1	23.3	24.9	25.2	25.0
Compensation	1.4	1.5	1.6	1.4	1.6	1.6
Investment	8.1	9.0	8.7	8.1	8.6	8.6
Free Cash Flow	4.0	4.4	4.7	4.1	4.7	3.9
Payout Rate	37.9	37.1	40.1	37.1	39.0	36.4
Entry/Exit Rate	1.2	0.8	0.1	1.8	0.4	0.4
Refinancing Rate	16.4	13.8	18.0	13.8	15.9	14.1
Refinancing Size	14.4	18.5	17.5	19.5	20.7	16.9
Distortions						
Agency (a)	2.54	2.64	1.66	2.92	2.14	1.83
Financial (f)	0.27	0.25	0.33	0.26	0.35	0.29
Ratio (a/f)	9.52	10.71	5.10	11.36	6.13	6.28
Relative to (1)	1.00	1.13	0.54	1.19	0.64	0.66

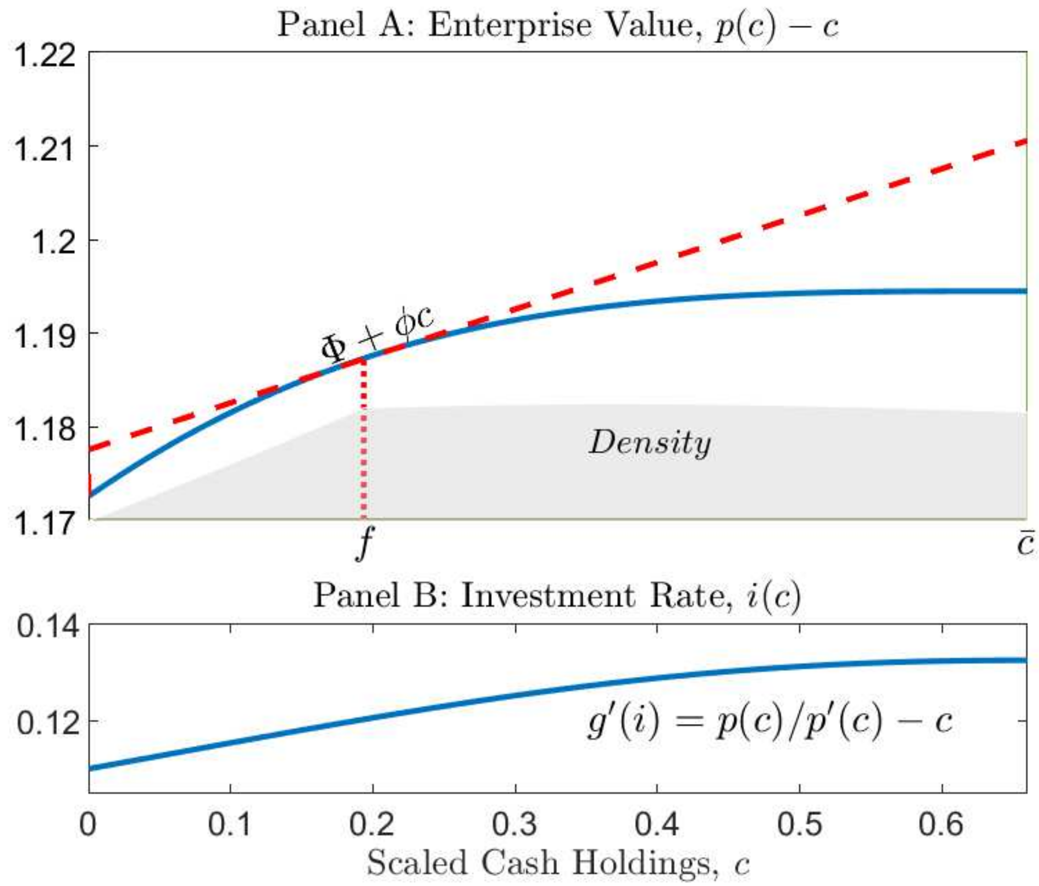


Figure 2.1: Cash Management with No Agency Conflict

Note: This figure depicts the solution to the cash management problem with no agency conflict under the parameters tabulated in Table 2.1 with the exception of setting $(\mu, z) = (0.16, 0.02)$. The domain for both panels is scaled cash holdings, $c = C/K \in [0, \bar{c}]$, where $\bar{c}$ is the payout boundary. In Panel A, the blue curve is the firm's enterprise value, $p(c) - c$. The red dashed line depicts marginal financing costs, $\Phi + \phi c$, and its tangency with enterprise value determines the refinancing size, f. The stationary density is unscaled to the vertical axis and shown in gray. Panel B plots the investment rate and its decision equation.

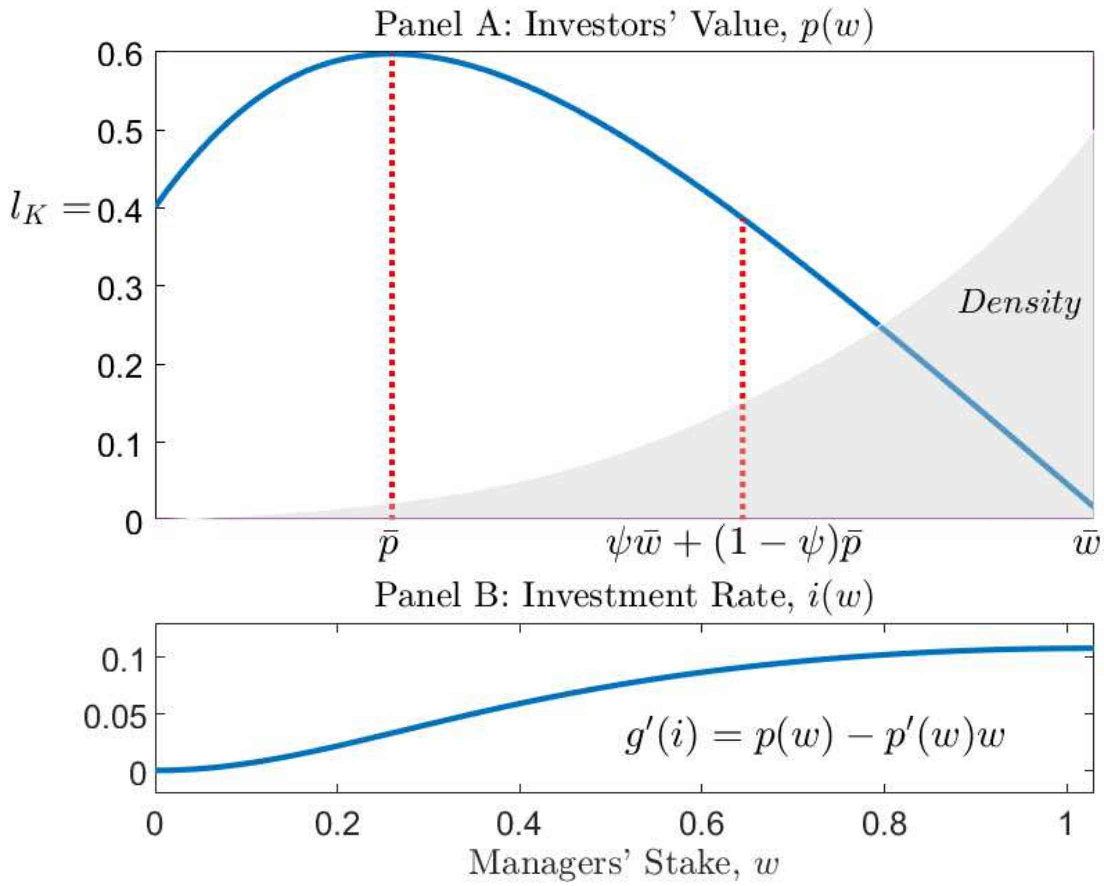


Figure 2.2: Agency Problem with Costless Refinancing

Note: This figure plots the solution to the agency problem with costless refinancing under the parameters tabulated in Table 2.1 with the exception of setting $(\mu, z) = (0.15, 0.02)$. The domain for both panels is managers' stake (scaled continuation payoff), $w = W/K \in [0, \overline{w}]$, where $\overline{w}$ is management's payment boundary. In Panel A, the blue curve is investors' scaled value function. At the termination boundary, $w = 0$, investors recover l_K, and a new firm is drawn with initial payoff $w_0 = \psi\overline{w} + (1-\psi)\overline{p}$, where $\overline{p} = \text{argmax}_w\, p(w)$. The stationary density is unscaled to the vertical axis and shown in gray. Panel B plots the investment rate and its decision equation.

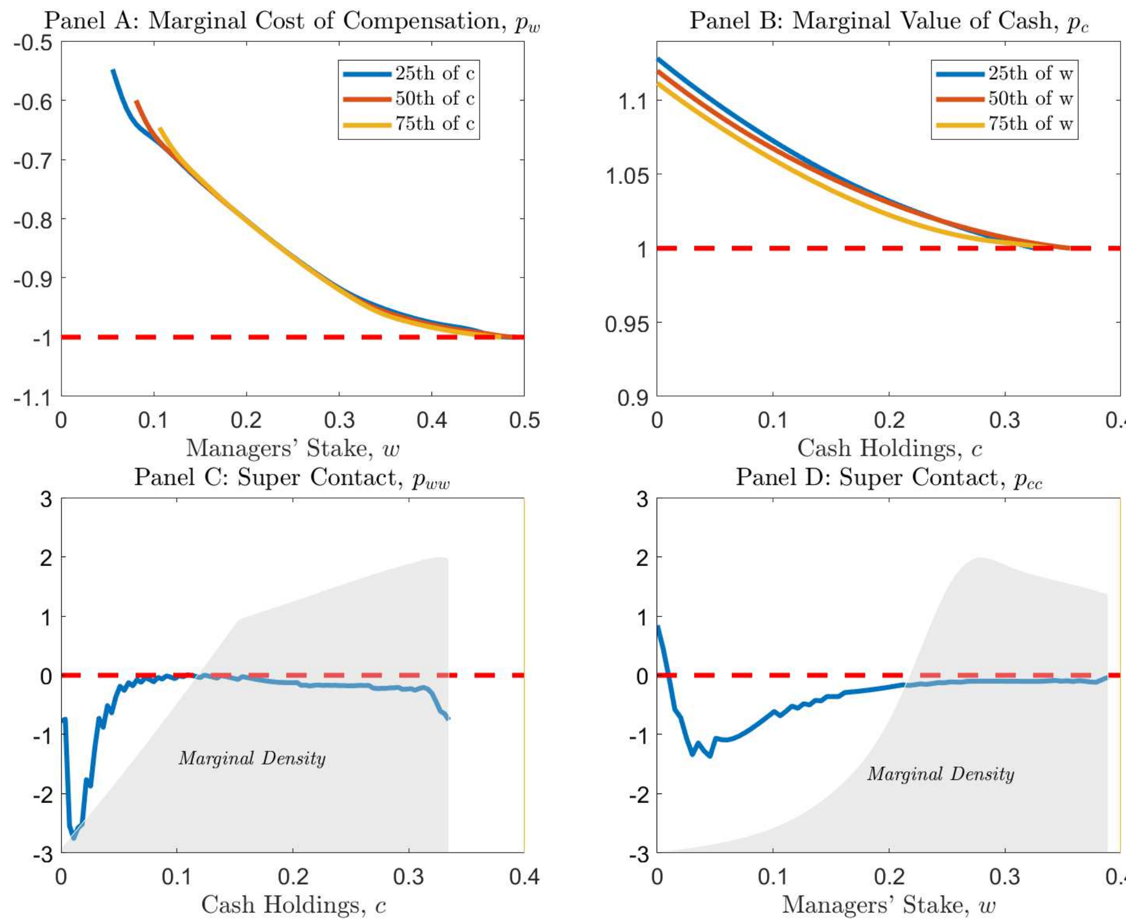


Figure 2.3: Accuracy of Model Solution

Note: This figure shows the accuracy of the complete model solution under the parameters tabulated in Table 2.1. Panel A plots the first derivatives of investors' scaled value function with respect to managers' stake (scaled continuation payoff), $w = W/K$, for percentiles of the marginal distribution of scaled cash holdings, $c = C/K$. Panel B plots the first derivatives of investors' scaled value function with respect to scaled cash holdings for percentiles of the marginal distribution of managers' stake. Panel C plots the super contact condition for the payment boundary, the second derivative of $p(c, \overline{w}(c))$ with respect to w for each value of c and Panel D plots the super contact condition for the payout boundary, the second derivative of $p(\overline{c}(w), w)$ with respect to c for each value of w. The marginal densities of scaled cash holdings and managers' stake are unscaled to the vertical axes and shown in gray.

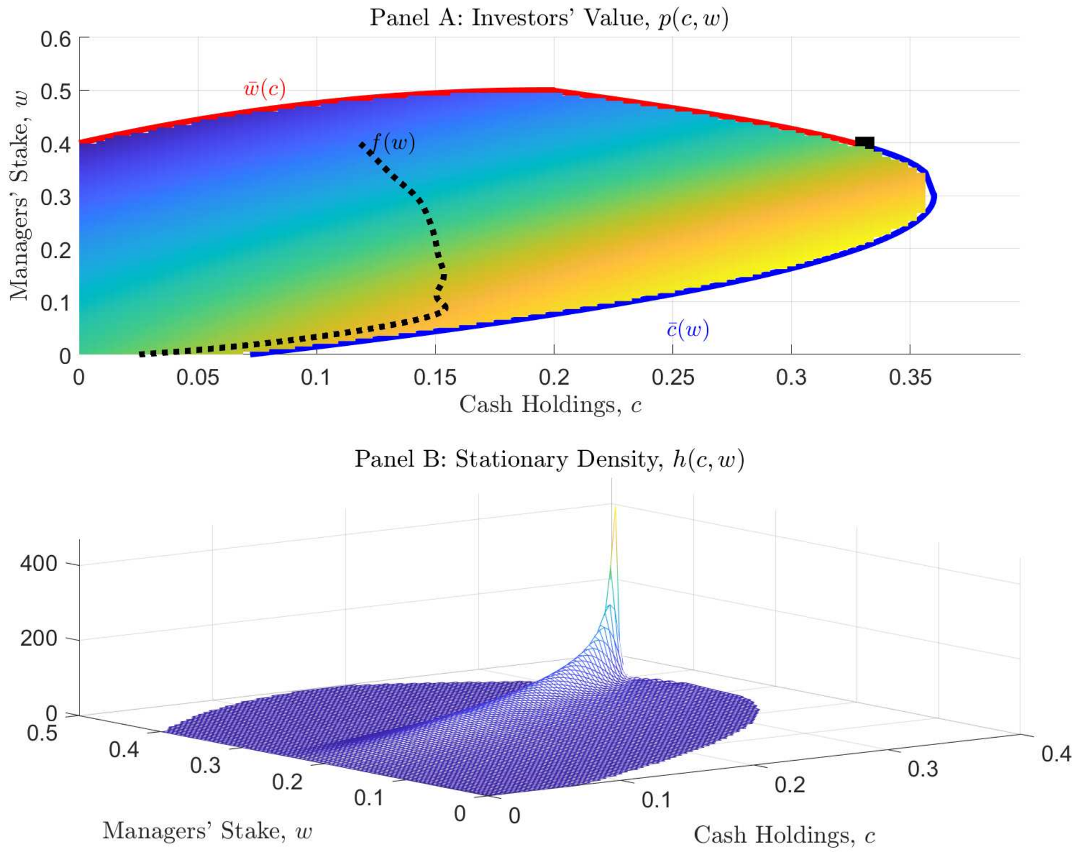


Figure 2.4: Properties of Model Solution

Note: This figure summarizes properties of the model solution under the parameters tabulated in Table 2.1. Panel A plots investors' scaled value function from above, bounded by the non-rectangular state space over managers' stake, $w = W/K$, and cash holdings, $c = C/K$. The black dotted line is the size of refinancing conditional on managers' stake, $f(w)$. The red line is the payment boundary, $\overline{w}(c)$, and the blue line the payout boundary, $\bar{c}(w)$. These boundaries intersect at the joint upper boundary, $(\bar{c}(w), \overline{w}(c))$, marked by the black square. Panel B plots the stationary density. In both panels a brighter color represents a higher value.

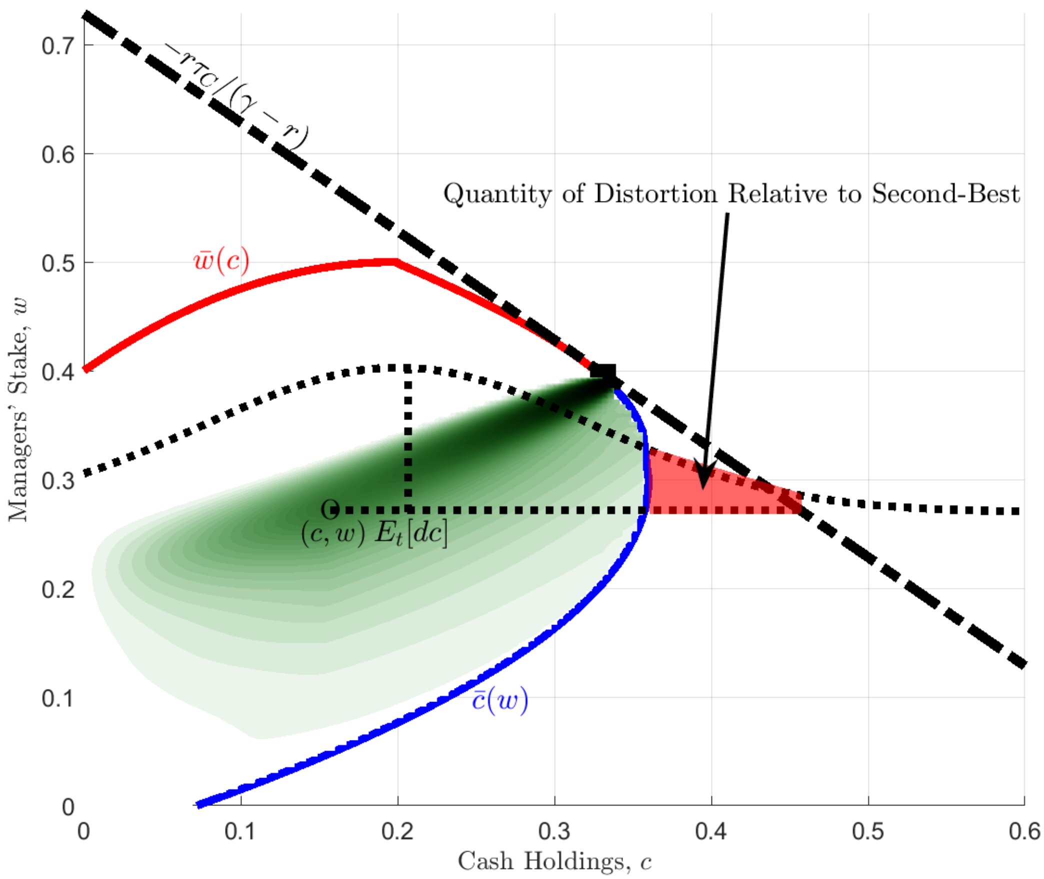


Figure 2.5: Illustration of Measurement of Distortions

Note: This figure illustrates the calculation used to measure the agency distortion relative to second-best; an analogous calculation holds for the financial distortion. It plots the stationary density in green from above, bounded by the state space over managers' stake, $w = W/K$, and scaled cash holdings, $c = C/K$. A darker green represents a higher value. The red line is the payment boundary, $\overline{w}(c)$, and the blue line the payout boundary, $\overline{c}(w)$. These boundaries intersect at the joint upper boundary, $(\overline{c}(w), \overline{w}(c))$, marked by the black square. The second-best frontier, the dash-dotted line, touches the joint upper boundary and has slope $dw/dc = -r\tau_C/(\gamma - r)$. Next, the point (c,w) marked by an O represents the density of firms at that point. The probability density of cash holdings over the next year point starting from O is normal with mean $\mathbb{E}_t[dc]$ that is marked on the figure and standard deviation $(1 - \tau_Y)\sigma$. This normal density is over the support denoted by the horizontal dotted line extending from O out to the second-best frontier, but is displayed with a tilt for the reader. The measurement for point (c,w) is the expected value over the interval starting at $\overline{c}(w)$ and ending at the frontier $\mathcal{F}^{-1}(w)$.

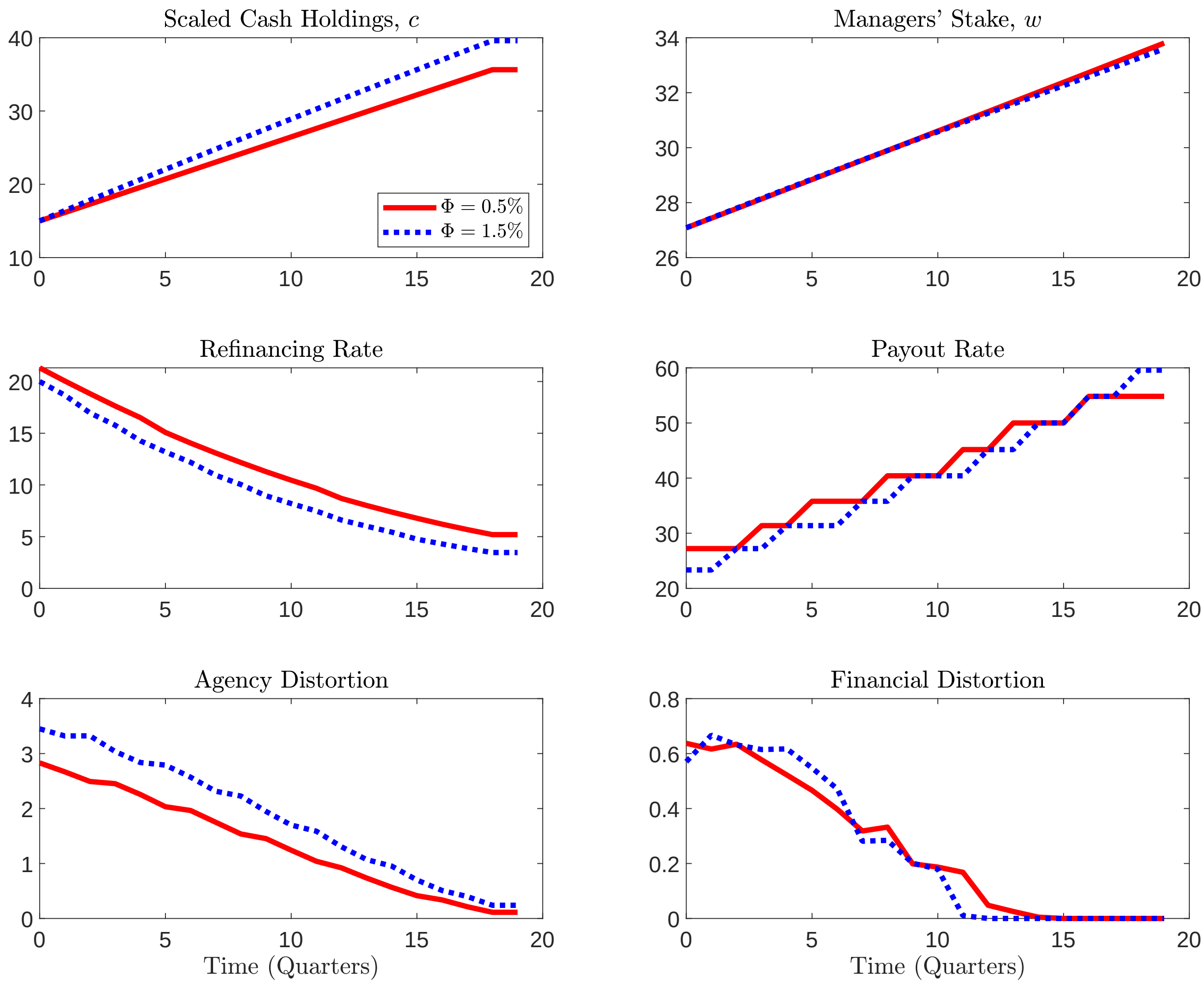


Figure 2.6: Impulse Dynamics Induced by Market Discipline

Note: This figure plots the impulse response function of several firm policies of the same firm initialized at (c, w) but contrasted across two economies that differ by fixed costs of refinance: $\Phi = 0.5$ or 1.5 percent. All variables are in percent. We initialize scaled cash holdings $c = C/K$ near the average refinancing amount in the data, $c_0 = 0.15$, and managers' stake at $w = (W/K)/\gamma = 0.013/\gamma$, where 1.3 percent is the average size of compensation in the data. Dynamics follow from the drifts of (2.38) and (2.39). Refinancing and payout statistics and agency and financial distortions are from (2.46), (2.47), (2.48) and (2.49), respectively.

Chapter 3

Debt Dynamics without Commitment in the Presence of Fixed Issuance Costs

3.1 Introduction

A large and growing literature in corporate finance investigates a firm's optimal dynamic capital structure policy through the lens of a tradeoff between the cost of financial distress and the benefit of tax deductibility of interest payments. In such a setting, shareholders may have incentive to issue additional debt over time both to increase tax benefits and to extract wealth from existing debtholders. Anticipating this incentive, creditors conjecture the firm's future debt issuance policy when pricing current debt. Thus, a firm's ability to commit to a future debt issuance policy may significantly impact its ability to extract tax benefits.

In this paper, we investigate a firm's optimal dynamic debt policy when shareholders cannot commit to future rebalancing policies and are subject to debt restructuring costs. Specifically, at any time, by paying a fixed cost, a firm can either issue additional debt or repurchase outstanding debt at market prices. In contrast to much of the existing literature, we do not assume that the firm *must* repay all debt outstanding prior to issuing new debt.[1] As emphasized by DeMarzo and He (2021), this counterfactual assumption is not innocuous. Indeed, we show that the modeling of non-callable debt in the presence of fixed restructuring costs significantly

[1] See, e.g., Fischer, Heinkel, and Zechner (1989), Leland (1998), Goldstein, Ju, and Leland (2001), Strebulaev (2007), Danis, Rettl, and Whited (2014), Hugonnier, Malamud, and Morellec (2015), and Dangl and Zechner (2020).

complicates the analysis, and has important implications for the existence of equilibria.

We first show that, in a Markov Perfect Equilibrium (MPE), shareholders never voluntarily reduce debt outstanding. This result extends the "leverage ratchet effect" of Admati, DeMarzo, Hellwig, and Pfleiderer (2018) to an economy with fixed restructuring costs and allows us to restrict the search for equilibria (but not the set of off-equilibrium deviations) to the class of Markovian strategies that involve only debt issuances. Within this class, we focus on policies of the "barrier" type specified in terms of (i) a single state variable, namely, the firm's income-to-debt ratio y, (ii) a default barrier y_b and a restructuring barrier y_u that characterize an inaction region (y_b, y_u), and (iii) a function $\mathcal{Y}(y)$ defined on the restructuring region $y \geq y_u$ that characterizes the amount of new debt that is issued to bring the firm's income-to-debt ratio back to an optimal target inside the inaction region.

We derive necessary and sufficient conditions for the existence of an MPE in barrier strategies. The interaction between issuance costs and debt maturity determines whether an equilibrium exists. More specifically, for issuance costs greater than some analytically-determined constant, the firm finds it optimal to never issue debt regardless of maturity, implying that our model reduces to that of Leland (1994). For issuance costs below this threshold, we identify three regions depending upon debt maturity. For sufficiently short maturity, we again find that the firm chooses not to issue debt, as shorter maturities imply higher rollover costs. As we increase debt maturity above an analytically-determined threshold, we identify a region in which shareholders extract positive tax benefits, with magnitude that varies with maturity, in spite of being unable to commit to an issuance strategy. Finally, as debt maturity increases further, we identify a region in which an MPE no longer exists. This is a surprising result because, in this region, rollover costs are relatively low, and thus significant tax benefits are available for firms that can commit to a dynamic debt policy. The lack of an equilibrium in this region is thus due to the myopic nature inherent in no-commitment policies. The intuition for the existence of this region is as follows: low rollover costs induce the manager to issue debt aggressively, i.e., to lower the restructuring boundary y_u. This in turn reduces the price creditors are willing to pay for the new debt issuance. But this lower price induces the manager to increase the location of the default boundary y_b, leading to a "vicious circle" of creditors offering lower debt prices and managers increasing y_b.

In order to quantify the loss of shareholder value due to their inability to commit to a particular strategy, we also explore the implications of our model for the case in which shareholders can commit to a future debt issuance policy. In particular, we identify a *global-optimal* strategy that maximizes equity value across all states of nature, subject to limited liability. Interestingly, if our model is calibrated to match typical empirical estimates of five-year maturity and one-percent debt issuance costs, we find that the tax benefits to debt in the MPE are only slightly lower than those of the global-optimal policy with commitment. We therefore conclude that, in the presence of realistic debt issuance costs, shareholders' inability to commit to a future debt issuance policy may have only a small impact on the tax benefits to debt that a firm can extract. Furthermore, we show that when the issuance cost parameter approaches zero, shareholders can extract close to 100% of the claim to the firm's cash flow for both cases with and without commitment. Thus, even for a case of vanishing, but positive, issuance costs, we see that not being able to commit to a dynamic capital structure policy may have minimal impact on a firm's ability to extract tax benefits to debt.

While our main focus is on MPEs, the global optimality of the policy with commitment allows us to also explore a possible equilibrium outside of the Markovian class. The optimal policy with commitment generates the highest equity values across all states of nature. Hence, shareholders do not have incentive to ever deviate if the punishment is that, following a deviation, debt is always priced according to the no-commitment equilibrium. Debtholders can credibly threaten such a punishment: Because they observe the size of the debt issuance and pay fair value for their claim, they are indifferent to the firm's restructuring policy. Under this "grim-trigger" punishment, the global optimal policy with commitment is therefore subgame perfect.

To our knowledge, this is the first paper that provides a formal characterization of barrier-strategy MPEs when debt issuance is subject to a fixed cost, and firms are not forced to retire outstanding debt prior to issuing new debt. The main challenge of this problem comes from the analysis of off-equilibrium deviations. When the income-to-debt ratio is outside the inaction region, that is, $y \geq y_u$, shareholders rebalance towards a target $\mathcal{Y}(y)$ that depends on the value of the state variable y. We provide a verification argument for the existence of an MPE and derive necessary and sufficient conditions for debt issuance to indeed be optimal in the restructuring region $y \geq y_u$. In contrast, when a firm is forced to retire all outstanding debt prior to issuing new debt, the issuance decision is always made with zero debt outstanding. Hence, when debt must be recalled prior to restructuring, the function $\mathcal{Y}(y)$ for all $y \geq y_u$ reduces to a single point $\mathcal{Y}(y = \infty)$, which greatly simplifies the analysis.

Our paper builds upon the quickly evolving literature that examines optimal dynamic capital structure decisions of firms. There are only a few tractable frameworks in this literature, due to the difficulty of valuing assets in an economy in which current prices depend on the firm's future debt issuance policy.[2] Most relevant to our work is DeMarzo and He (2021), who investigate

[2] See, for example, He and Milbradt (2016), Admati, DeMarzo, Hellwig, and Pfleiderer (2018), DeMarzo, He, and Tourre (2018), DeMarzo (2019), Benzoni, Garlappi, and Goldstein (2020), the literatures on roll-over debt structure, e.g., Leland and Toft (1996), He and Xiong (2012a), Cheng and Milbradt (2012), Décamps and Villeneuve (2014), Della Seta, Morellec, and Zucchi (2020), as well as the literature on optimal maturity choice, e.g., Leland (1998), Brunnermeier and Yogo (2009), He and Xiong (2012b), Chen, Xu, and Yang (2012), Brunnermeier and Oehmke (2013), Diamond and He (2014), He and Milbradt (2014), Abel (2016).

leverage dynamics without commitment in the absence of rebalancing costs. They show that the unique MPE is characterized by a locally deterministic process in which new debt is issued in all states of nature, even when the firm is near default. Due to this aggressive policy, there are no tax benefits to debt regardless of maturity. In contrast, in our model, shareholders can extract positive tax benefits, the size of which depends on maturity. Moreover, tax benefits tend to increase as issuance costs decline. Thus, our model does not converge to that of DeMarzo and He (2021) even in the presence of arbitrarily small issuance costs, in spite of the fact that our model *is* identical to theirs when the issuance cost parameter is set to zero at the outset. The intuition for this result is the following: regardless of how small the debt issuance parameter is, so long as it is positive, optimal debt issuance is not locally deterministic, because that would imply an infinite accumulation of issuance costs in a continuous time framework. Hence, in our model the equilibrium policy is always characterized by a region of inaction defined by a lower default boundary and an upper debt-issuance boundary.

Another related paper is Malenko and Tsoy (2020), who consider a model of debt issuance and repurchases without restructuring costs. They focus on time-consistent barrier policies that satisfy a credibility constraint when EBIT follows a jump-diffusion process, which in turn identifies a subgame-perfect equilibrium outside of the Markov class. Aside from the inclusion of jumps and the absence of issuance costs, their results are analogous to our identification of equilibria with commitment sustained by a “grim-trigger” punishment, in which any deviation would lead to debt being priced as in the no-commitment economy.

A vast literature has studied dynamic capital structure choice in the presence of issuance costs.[3] Two papers in this literature are closely related to our work. First, Goldstein, Ju, and Leland (2001) show that the firm can extract positive tax benefit from debt in a model with perpetual callable debt, proportional issuance costs, and commitment. We extend their analysis to the case of non-callable debt, finite maturity, fixed issuance costs, and no-commitment, and explore the interplay between maturity and issuance costs. Second, Dangl and Zechner (2020) also study the interplay between issuance costs and maturity when the firm chooses how to refinance expiring debt. Similar to our findings, Dangl and Zechner (2020) highlight a tradeoff between the costs of higher rollover frequencies and the benefits of increased flexibility associated with shorter maturity debt. Unlike Dangl and Zechner (2020), we focus on non-callable debt that does not have to be retired prior to additional debt being issued. Moreover, we formally derive necessary and sufficient conditions for the existence of an MPE in barrier strategies. Finally, our setting allows us to study debt policies with fixed issuance costs in an economy

[3] See, e.g., Kane, Marcus, and McDonald (1984, 1985), Fischer, Heinkel, and Zechner (1989), Titman and Tsyplakov (2007), Strebulaev (2007), Morellec, Nikolov, and Schürhoff (2012), Hennessy and Whited (2007), Gomes and Schmid (2012), Bolton, Chen, and Wang (2011b), Hugonnier, Malamud, and Morellec (2015), Bolton, Wang, and Yang (2020), and many others.

that is otherwise identical to that of DeMarzo and He (2021) and, moreover, to quantify the difference in available tax benefits between models with and without commitment.

The predictions of our model are in line with the empirical literature that investigates the capital structure and maturity decisions of firms. The presence of fixed restructuring costs in our model predicts that firms issue debt in discrete (rather than continuous) amounts, consistent with observation. Moreover, our model generates both persistence in leverage and a negative correlation between profitability and leverage, consistent with, e.g., Titman and Wessels (1988) and Frank and Goyal (2014). Our model captures these features because, when firms are in the inaction region, higher profitability increases equity values while debt outstanding remains constant, leading to lower leverage, and vice-versa. Also consistent with our model's predictions are van Binsbergen, Graham, and Yang (2010), and Blouin, Core, and Guay (2010), who document that firms are able to extract tax benefits to debt. Moreover, our findings are consistent with Barclay and Smith (1995) and Stohs and Mauer (1996), who report that firms are not indifferent toward debt maturity choice. Fama and French (2002), Baker and Wurgler (2002), and Welch (2004) provide evidence that shocks to capital structures are persistent, and Leary and Roberts (2005) attribute this persistence to the presence of adjustment costs. Finally, Graham and Harvey (2001) report survey evidence that 45% of CFOs are concerned with the tax advantage of interest deductibility, suggesting that firms are not indifferent to capital structure choices.

The rest of the paper is as follows. In Section 3.2, we present the model, prove necessary and sufficient conditions for the existence of a barrier-strategy MPE, and derive a policy with commitment that maximizes equity values within the class of barrier strategies. In Section 3.3, we study the properties of the MPEs in comparison with the global optimal policy with commitment. Section 3.4 concludes. Appendix A contains the formal definition of strategies and equilibrium and Appendix B contains propositions and proofs. Additional supporting results are in the Online Appendix.

3.2 The Model

The firm. We consider an economy in which all agents are risk neutral and the discount rate $r > 0$ is exogenous. The representative firm is characterized by two state variables. The first is the exogenously specified EBIT process Y_t, with dynamics given by:

$$\frac{dY_t}{Y_t} = \mu\, dt + \sigma\, dW_t, \tag{3.1}$$

for some constants $\mu < r$ and $\sigma > 0$, where dW_t denotes increments of a standard Brownian motion. The value V_t of the claim to EBIT is:

$$V_t = \mathbb{E}_t\left[\int_t^{\infty} e^{-r(s-t)}\, Y_s\, ds\right] = \frac{Y_t}{r-\mu}. \tag{3.2}$$

Due to the linear relation between Y_t and V_t in equation (3.2), we can choose either as the exogenous state variable. We choose Y_t for consistency with the existing dynamic capital structure literature.

The second state variable is the date-t outstanding face value of debt F_t associated with a coupon rate c and inverse maturity ξ. During the interval $(t, t+dt)$, debtholders receive cash flows $(c+\xi)F_t\, dt$ as long as the firm operates. As in the benchmark case of DeMarzo and He (2021), we assume that recovery of bonds in default is zero.

The dynamics for the face value of debt are endogenously determined in that, by paying a fixed adjustment cost βY_t at any time, the firm can adjust its capital structure by retiring or issuing bonds at market prices. The adjustment cost is fixed in that it is independent of the size of debt adjustment. As a result of debt adjustments, the face value of the firm's debt evolves according to the process

$$dF_t = -\xi F_{t-}\, dt + A_t\, F_{t-}\, dN_t, \tag{3.3}$$

where $A_t = \left(\frac{F_t - F_{t-}}{F_{t-}}\right)$ is the fractional change in outstanding debt at a restructuring time t, and dN_t is a counting process which increases by unity each time debt is restructured. Both (A_t, dN_t) should be understood as policy decisions of the firm.

We look for equilibria in which all asset claims $H(Y_t, F_t)$ are homogeneous of degree one with respect to the state vector, i.e., $H(\lambda Y_t, \lambda F_t) = \lambda H(Y_t, F_t)$. By choosing $\lambda = 1/F_t$, we can express the value of any claim, scaled by the face value of debt, as a function of a single state variable, namely, the income-to-debt ratio:

$$y_t \equiv \frac{Y_t}{F_t}, \tag{3.4}$$

that is,

$$\frac{H(Y_t, F_t)}{F_t} = H\left(\frac{Y_t}{F_t}, 1\right) \equiv h(y_t). \tag{3.5}$$

Below, we investigate (reduced) Markov-Perfect Equilibria (MPE hereafter) in which the strategy followed by management is characterized solely in terms of the state variable y_t.

3.2.1 Characterization of Markov perfect equilibria

In our framework, a strategy is Markov if for each value of the state variable y_t, management makes one of three choices: (i) default, (ii) change the level of debt outstanding; or (iii) do neither, and simply service outstanding debt. Definition 4 in Appendix A formalizes the concept of MPE for a general class of default and debt restructuring strategies. The characterization of an MPE for such a large class of strategies is impractical. However, we are able to derive results that allow us to restrict the analysis to a tractable subset of strategies. In particular, in

Proposition 7 we show that in every MPE, firms never find it optimal to repurchase debt. This result, which extends the "leverage ratchet" effect (see, e.g., Admati, DeMarzo, Hellwig, and Pfleiderer (2018) and DeMarzo and He (2021)) to the case of fixed issuance costs, allows us to consider only debt *issuance* strategies. Furthermore, in Proposition 8 we show that in any MPE the equity value is a non-decreasing function of the income-to-debt ratio y_t that vanishes when y_t falls below a constant threshold y_b. These two results allow us to restrict the search of MPEs (but not the set of possible deviations) to the subset of debt issuance and default strategies characterized by a constant default threshold.

3.2.2 Markov perfect equilibria in barrier strategies

A class of strategies of particular interest within the set identified in Section 3.2.1 is the class of *barrier* strategies $\mathcal{B}$ that can be characterized by two parameters and one function. The two parameters are the default boundary (y_b) and the debt-issuance boundary (y_u), which satisfies the restriction $(y_u > y_b)$. These two parameters demarcate three regions. The region $y_t \leq y_b$ is the *default region*, where it is optimal for management to immediately default, rather than continue to service outstanding debt. The region $y_t \in (y_b, y_u)$ is the *inaction region*, where it is optimal for management to neither default nor change the level of outstanding debt. Finally, the region $y_t \geq y_u$ is the *restructuring region*, where it is optimal to issue debt in sufficient amounts so that the post-issuance income-to-debt ratio immediately returns to the inaction region. The function $\mathcal{Y}(y_t) : [y_u, \infty) \to (y_b, y_u)$ is the *restructuring function* associated with the parameters y_b and y_u. Following a restructuring, the income-to-debt ratio jumps from $y_{t-} \in [y_u, \infty)$ to $y_{t+} \equiv \mathcal{Y}(y_{t-}) \in (y_b, y_u)$.[4] An important special case is $\hat{y} \equiv \mathcal{Y}(y_u)$, which is the target income-to-debt ratio chosen at the upper boundary of the inaction region.

Note that, under a barrier strategy, if the state variable "begins" in the inaction region, then the entire *restructuring region* is inaccessible, except for y_u. Yet, as we discuss in Section 3.2.3, the existence of a barrier equilibrium relies on a verification argument for all off-equilibrium values $y_t \in (y_u, \infty)$. Applying Ito's lemma to equation (3.4), we find:

$$\begin{aligned} dy_t &= \frac{dY_t}{F_{t^-}} - \frac{Y_t}{F_{t^-}^2}\left(-\xi F_{t^-}\, dt\right) + \left(\frac{Y_t}{F_{t^-}(1+A_t)} - \frac{Y_t}{F_{t^-}}\right) \mathbf{1}_{\{y_{t^-} \geq y_u\}} \\ &= y_{t^-}\left[\, (\mu+\xi)\, dt + \sigma\, dW_t\right] + \underbrace{\left(\mathcal{Y}(y_{t^-}) - y_{t^-}\right)}_{\leq 0} \mathbf{1}_{\{y_{t^-} \geq y_u\}}, \end{aligned} \tag{3.6}$$

where, consistent with the focus on barrier strategies, we have replaced the counting process (dN_t) with an indicator function that equals one if and only if the current state vector is in the

[4] In Corollary 2 of the appendix, we prove that it is never optimal to jump either to the default region or to another location in the restructuring region.

restructuring region, and where we have used the relation $\mathcal{Y}(y_{t^-}) = \frac{y_{t-}}{1+A_t}$.

Debt valuation. Consider a representative debtholder who at date-t owns the entire debt claim with face value F_t. Due to the bond's exponential maturity, the face value of debt owned by this agent decays as $F_s = F_t\, e^{-\xi(s-t)}$. Hence, the amount of maturing principal paid out over the interval ds is $-dF_s = \xi F_t\, e^{-\xi(s-t)}\, ds$. Thus, the cash flow received by creditors over the interval ds, which is composed of both coupon and principal payments, is equal to $(c+\xi)F_t e^{-\xi(s-t)}\, ds$. The debt price per unit face value is therefore:[5]

$$p(y_t) = \mathbb{E}_t\left[\int_t^{\tau_b} e^{-(r+\xi)(s-t)}(c+\xi)\, ds\right], \tag{3.7}$$

where τ_b denotes the first time that the income-to-debt ratio y_t hits the default boundary y_b from above, that is, $\tau_b := \inf\{s > t : y_s \leq y_b\}$. Given the zero-recovery assumption, we thus identify the conditions:

$$p(y_t) \quad = \quad 0 \qquad \forall y_t \in (0, y_b]. \tag{3.8}$$

Similarly, debt issuance occurs when the current income-to-debt ratio is in the restructuring region $[y_u, \infty)$. Importantly, to preclude arbitrage opportunities, the bond price per unit face value of debt must not jump at the debt issuance time:

$$p(y_t) \quad = \quad p(\mathcal{Y}(y_t)) \qquad \forall y_t \in [y_u, \infty). \tag{3.9}$$

Standard arguments relying on Itô's lemma and the continuity of the bond price at the issuance boundary show that, for values of $y_t \in (y_b, y_u)$, the debt price in equation (3.7) is the unique solution $p(y)$ to the following ordinary differential equation:

$$0 = (c+\xi) - (r+\xi)\,p(y) + (\mu+\xi)y\,p'(y) + \frac{\sigma^2}{2}y^2\,p''(y), \quad y \in (y_b, y_u), \tag{3.10}$$

subject to the boundary conditions

$$\begin{aligned} p(y) &= 0 & 0 \leq y \leq y_b \qquad (3.11)\\ p(y) &= p(\mathcal{Y}(y)) & y \geq y_u. \qquad (3.12)\end{aligned}$$

[5] Note that the firm may issue additional debt after date-t, which impacts τ_b, but not the validity of equation (3.7).

The solution is

$$p(y_t) = \begin{cases} 0 & \text{if} \quad 0 \leq y_t \leq y_b \\ \left(\frac{c+\xi}{r+\xi}\right)\left(1 + M_p\, y_t^{\Theta} + N_p\, y_t^{\Pi}\right) & \text{if} \quad y_b < y_t < y_u \\ \left(\frac{c+\xi}{r+\xi}\right)\left(1 + M_p\, \mathcal{Y}(y_t)^{\Theta} + N_p\, \mathcal{Y}(y_t)^{\Pi}\right) & \text{if} \quad y_t \geq y_u \end{cases}, \tag{3.13}$$

with $M_p \equiv \frac{y_u^{\Pi} - \hat{y}^{\Pi}}{y_b^{\Pi}\left(y_u^{\Theta} - \hat{y}^{\Theta}\right) + y_b^{\Theta}\left(\hat{y}^{\Pi} - y_u^{\Pi}\right)} \leq 0$ and $N_p \equiv \frac{\hat{y}^{\Theta} - y_u^{\Theta}}{y_b^{\Pi}\left(y_u^{\Theta} - \hat{y}^{\Theta}\right) + y_b^{\Theta}\left(\hat{y}^{\Pi} - y_u^{\Pi}\right)} \leq 0$, and where the exponents (Θ, Π)

$$\Theta = \frac{\frac{\sigma^2}{2} - (\mu + \xi) + \sqrt{\left(\frac{\sigma^2}{2} - (\mu + \xi)\right)^2 + 2(r + \xi)\sigma^2}}{\sigma^2} > 1 \tag{3.14}$$

$$\Pi = \frac{\frac{\sigma^2}{2} - (\mu + \xi) - \sqrt{\left(\frac{\sigma^2}{2} - (\mu + \xi)\right)^2 + 2(r + \xi)\sigma^2}}{\sigma^2} < 0 \tag{3.15}$$

are the two solutions $\lambda_{\pm}$ to the quadratic equation

$$-(r + \xi) + (\mu + \xi)\lambda + \frac{\sigma^2}{2}\lambda(\lambda - 1) = 0. \tag{3.16}$$

Because $(M_p, N_p) \leq 0$ and $y_b < \hat{y} < y_u$, if follows that the bond price function is strictly concave in the inaction region (y_b, y_u), with $p'(y_b) > 0$, $p'(\hat{y}) \geq 0$, and $p'(y_u) \leq 0$.

Equity valuation without commitment. Because the decision to default or adjust the level of outstanding debt is made when debt has been issued previously, shareholders may have an incentive to deviate from the policy conjectured by creditors. Absent commitment, this implies that creditors will lend money to the firm only if the debt contract is *incentive compatible* in that shareholders would never want to deviate from the default and issuance strategy that creditors use to price the bonds at issuance. Here we formalize the notions of incentive compatibility and Markov Perfect Equilibria.

The firm's EBIT Y_t is subject to a corporate tax rate $\tau \in [0, 1)$ and coupon payments are tax deductible. Hence, the instantaneous cash flow to equity holder, i.e., the dividend, is given by

$$\delta(F_t, Y_t) \equiv (1 - \tau)Y_t - (c(1 - \tau) + \xi)F_t. \tag{3.17}$$

If bondholders conjecture that the firm will use a barrier strategy $\mathbf{a} \in \mathcal{B}$ but shareholders instead use a different barrier strategy $\mathbf{s} \in \mathcal{B}$, then the value of equity is

$$e(y_t|(\mathbf{s},\mathbf{a})) = \mathbb{E}_t\left[\int_t^{\min(\tau_b(\mathbf{s}),\tau_u(\mathbf{s}))} e^{-(r+\xi)(s-t)}\,\delta(y_s)\,ds + \mathbf{1}_{(\tau_u(\mathbf{s})<\tau_b(\mathbf{s}))}\,e^{-(r+\xi)(\tau_u(\mathbf{s})-t)}e(y_u|(\mathbf{s},\mathbf{a}))\right], \tag{3.18}$$

where we have defined $\delta(y_t) \equiv \delta(1, y_t)$.

Definition 3. *A Markov Perfect Equilibrium (MPE) in barrier strategies is a strategy* $\mathbf{a} \in \mathcal{B}$ *such that*

$$e(y_t|(\mathbf{a},\mathbf{a})) = \sup_{\mathbf{s}\in\mathcal{B}} e(y_t|(\mathbf{s},\mathbf{a})), \quad t \geq 0. \tag{3.19}$$

The definition formalizes the fact that in an MPE shareholders have no incentive to deviate from the creditors' conjectured strategy.

Equation (3.18) holds for the inaction region (y_b, y_u) associated to a particular barrier strategy $\mathbf{a} \in \mathcal{B}$. Due to absolute priority, the equity claim is zero in the default region, that is,

$$e(y_t) \quad = \quad 0 \qquad \forall y_t \in (0, y_b]. \tag{3.20}$$

In the restructuring region $[y_u, \infty)$, to preclude arbitrage, the value of the equity claim just prior to debt restructuring must equal the equity claim just after restructuring plus any cash flows received from debt issuance. Denoting by $E(F_t, Y_t)$ the equity value in the unscaled economy, for values of the state variables F_t and Y_t such that $y_t \equiv \frac{Y_t}{F_t} \in [y_u, \infty)$, we have that

$$E(F_t, Y_t) = \sup_{\hat{F}_t \geq 0}\left\{E(\hat{F}_t, Y_t) + (\hat{F}_t - F_t)\,p\Big(Y_t/\hat{F}_t\Big) - \beta Y_t\right\}. \tag{3.21}$$

The second term on the right-hand side is the cash flow to equity from the debt issuance of size $(\hat{F}_t - F_t)$, which is priced by creditors at the *post-adjustment* price $p\Big(Y_t/\hat{F}_t\Big)$. Using the fact that $E(\hat{F}_t, Y_t) = \hat{F}_t e\left(\frac{Y_t}{\hat{F}_t}\right)$, and denoting $z_t \equiv \frac{Y_t}{\hat{F}_t}$ as the optimal income-to-debt target for each value of $y_t = \left(\frac{Y_t}{F_t}\right)$, we obtain the scaled version of the continuation value in equation (3.21):

$$e(y_t) \;=\; \frac{E(F_t, Y_t)}{F_t} \;=\; \sup_{z_t \geq 0}\left\{\left(\frac{y_t}{z_t}\right)e(z_t) + \left(\frac{y_t}{z_t} - 1\right)p(z_t) - \beta y_t\right\}. \tag{3.22}$$

The maximization in equation (3.22) defines the restructuring function $\mathcal{Y}(y_t)$ associated with the parameters y_b and y_u. Specifically, we have

$$\mathcal{Y}(y) = \operatorname*{argmax}_{z\in(y_b,y_u)}\left\{\frac{y}{z}e(z) + \left(\frac{y}{z} - 1\right)p(z) - \beta y\right\}, \quad y \geq y_u. \tag{3.23}$$

For the special case $(y = y_u)$, we denote $\hat{y} \equiv \mathcal{Y}(y_u)$.

Standard results show that the equity value in equation (3.18) is the solution to the following ordinary differential equation

$$0 = \delta(y) - (r+\xi)\, e(y) + (\mu+\xi) y\, e'(y) + \frac{\sigma^2}{2} y_t^2\, e''(y) \qquad \forall y \in (y_b, y_u), \tag{3.24}$$

subject to the boundary conditions

$$\begin{aligned} e(y) &= 0 & 0 \le y \le y_b \qquad (3.25)\\ e(y) &= \frac{y}{\mathcal{Y}(y)} e(\mathcal{Y}(y)) + \left(\frac{y}{\mathcal{Y}(y)} - 1\right) p(\mathcal{Y}(y)) - \beta y & y \ge y_u. \qquad (3.26)\end{aligned}$$

The solution to this problem is

$$e(y_t) = \begin{cases} 0 & \text{if} \quad 0 \le y_t \le y_b \\ \hat{e}(y_t) + M_e\, y_t^{\Theta} + N_e\, y_t^{\Pi} & \text{if} \quad y_b < y_t < y_u \\ \hat{e}(\mathcal{Y}(y_t)) + M_e\, \mathcal{Y}(y_t)^{\Theta} + N_e\, \mathcal{Y}(y_t)^{\Pi} & \text{if} \quad y_t \ge y_u, \end{cases} \tag{3.27}$$

with $\hat{e}(y)$ denoting the levered claim to EBIT,

$$\hat{e}(y_t) = -\frac{c(1-\tau)+\xi}{r+\xi} + \left(\frac{1-\tau}{r-\mu}\right) y_t, \tag{3.28}$$

and where the constants M_e, and N_e are the unique solutions to the value-matching conditions

$$\begin{aligned} e(y_b) &= 0 & (3.29)\\ e(y_u) &= \frac{y_u}{\hat{y}} e(\hat{y}) + \left(\frac{y_u}{\hat{y}} - 1\right) p(\hat{y}) - \beta y_u, \quad \text{with } \hat{y} \equiv \mathcal{Y}(y_u). & (3.30)\end{aligned}$$

Equations (3.27)-(3.30) hold for any given $(y_b,\, y_u)$. Proposition 9 in Appendix B shows that a necessary condition for a barrier strategy $\mathbf{a} \in \mathcal{B}$ to be an MPE is that the equity value function $e(y_t)$ satisfies smooth-pasting conditions at the default and restructuring boundaries y_b and y_u, that is,

$$\begin{aligned} e'(y_b) &= 0 & (3.31)\\ e'(y_u) &= \frac{e(\hat{y}) + p(\hat{y})}{\hat{y}} - \beta. & (3.32)\end{aligned}$$

3.2.3 Existence of Markov perfect equilibria in barrier strategies

Up to this point, most of our focus has been on the inaction region (y_b, y_u). This focus has allowed us to construct a *candidate* MPE within the class of barrier strategies formally characterized by

Proposition 9. However, in order to insure that such a candidate strategy is indeed an MPE, we need to verify that shareholders have no incentive to deviate from the strategy conjectured by creditors. The following proposition provides a verification argument that identifies necessary and sufficient conditions for an MPE to exist.

Proposition 4 (Verification argument for MPE existence)**.** *Consider a barrier strategy* $\mathbf{a} = (y_b, y_u, \mathcal{Y}(y))$ *that satisfies the necessary conditions of Proposition 9 and let*

$$\Phi(y, z|\mathbf{a}) \equiv \frac{y}{z} e(z|\mathbf{a}) + \left(\frac{y}{z} - 1\right) p(z|\mathbf{a}) - \beta y, \quad y, z \geq 0. \tag{3.33}$$

Then such a strategy $\mathbf{a}$ *satisfies*

$$\sup_{z \in [y_b, y_u]} \Phi(y, z|\mathbf{a}) = \sup_{z \geq 0} \Phi(y, z|\mathbf{a}), \quad y \geq 0, \tag{3.34}$$

and constitutes an MPE if and only if

$$e(y|\mathbf{a}) \geq \Phi(y, z|\mathbf{a}), \quad (y, z) \in [0, y_u]^2, \tag{3.35}$$

and the following condition holds for all $y \geq y_u$*:*

$$\delta(y) - (r - \mu)\, e(y|\mathbf{a}) + (\mu + \xi)\, p(y|\mathbf{a}) + \frac{1}{2}\sigma^2 y^2\, p'(y|\mathbf{a}) \quad \leq \quad 0. \tag{3.36}$$

In the case of $y \to \infty$*, condition* (3.36) *simplifies to*

$$\delta(y) - (r - \mu) e(y|\mathbf{a}) \leq 0. \tag{3.37}$$

Equation (3.34) is a consequence of Corollary 2, which proves that it is never optimal to restructure in a way such that the post debt-issuance income-to-debt ratio falls either in the default region, $y < y_b$, or in the restructuring region $y > y_u$. The condition in equation (3.35) guarantees that it is not optimal to issue debt when the income to debt ratio y is inside the inaction region (y_b, y_u). Equations (3.36)–(3.37) are off-equilibrium conditions guaranteeing that it is optimal to issue debt when the income to debt ratio falls in the restructuring region, $y \geq y_u$.

To provide intuition for the off-equilibrium conditions (3.36)–(3.37), we note that for a policy to be optimal, shareholders cannot be better off following any other strategy. In this regard, we compare two strategies. The first strategy is to follow the proposed optimal policy by immediately issuing an optimal amount of debt as described in equation (3.21) above. The second strategy is to wait a period dt (in turn, receiving the cash flows owed to shareholders), and then issue an optimal amount of debt according to equation (3.21). Hence, for a proposed strategy to be optimal, for all values of $(Y_t,\, F_t)$ such that $y_t \equiv \frac{Y_t}{F_t} \in [y_u, \infty)$, it must be that:

$$E(F_t, Y_t) \quad \geq \quad \delta(F_t, Y_t)\, dt + e^{-r\, dt}\, \mathbb{E}_t \left[E(F_t + dF_t, Y_t + dY_t)\right]. \tag{3.38}$$

Using Itô's lemma and recalling that $E(F_t, Y_t) = F_t e(y_t)$ and $\delta(F_t, Y_t) = F_t \delta(y_t)$, equation (3.38) simplifies to:

$$0 \geq \delta(y_t) - (r+\xi)\, e(y_t) + (\mu+\xi) y\, e'(y_t) + \frac{\sigma^2}{2} y_t^2\, e''(y_t). \tag{3.39}$$

In the proof of Proposition 4 we show that equation (3.39) is equivalent to equation (3.36). Equation (3.38) describes the problem of an agent whose only available gamble is associated with an expected loss, and whose only decision is to choose when to stop playing. Thus, we refer to these alternative, but equivalent, formulations of the off-equilibrium condition as a "supermartingale condition."

Finally, equation (3.37) implies that the claim to equity for an unlevered firm is higher if shareholders issue debt according to the MPE policy rather than remain unlevered forever: $e(y_t) \geq \delta(y_t)/(r-\mu) = (1-\tau)v_t$.

3.2.4 Commitment equilibria in barrier strategies

So far we have focused on Markov-perfect equilibria in which shareholders do not have the ability to commit to a future debt issuance policy. To assess how valuable the ability to commit is to shareholders, here we investigate benchmark models in which shareholders can commit to a future restructuring policy.[6] Specifically, we generalize two models that have been widely studied in the literature. First, we investigate a version of the Leland (1994) model in which the firm can commit to never restructure its debt, i.e., $y_u \to \infty$. Second, we consider a version of the Goldstein, Ju, and Leland (2001) model in which shareholders commit to an optimal restructuring threshold y_u and target $\hat{y}$. In both cases, we generalize the original model by allowing for finite maturity debt.

Generalized Leland model

We specify the EBIT process Y_t as in equation (3.1). Similar to equation (3.3), we specify existing debt with a constant coupon rate c and constant amortization rate ξ. However, here we assume that no additional debt can be issued in the future (i.e., $A_t = 0$):

$$dF_t = -\xi F_{t-}\, dt. \tag{3.40}$$

[6] It is reasonable to assume that firms can never credibly commit to a default boundary y_b prior to issuing debt. However, as we discuss in more detail below, debt issuance is potentially a repeated game, hence there is the possibility that a policy with commitment to a restructuring boundary is a subgame perfect equilibrium.

At date 0, the firm has EBIT Y_0 and no debt outstanding, i.e., $F_0 = 0$. If it never issues debt, equity value is:

$$E_{\text{no issue}}(Y_0, 0) \;=\; \left(\frac{1-\tau}{r-\mu}\right) Y_0. \tag{3.41}$$

Instead, if it decides to make a one-time debt issuance, it will choose this level optimally:

$$E_{\text{issue}}(Y_0, 0) \;=\; \max_{F \geq 0} \left\{ E(Y, F) + P(Y, F) - \beta Y_0 \right\}. \tag{3.42}$$

Denoting $y \equiv Y/F$, we rewrite this optimization as:

$$\frac{E_{\text{issue}}(Y_0, 0)}{Y_0} \;=\; \max_{\hat{y}_{\text{Leland}}} \left\{ \frac{e(\hat{y}_{\text{Leland}}) + p(\hat{y}_{\text{Leland}})}{\hat{y}_{\text{Leland}}} \right\} - \beta. \tag{3.43}$$

Note that when $y_u = \infty$, there is no difference between models with and without commitment. Also note that the value of equity is the maximum:

$$E(Y_0, 0) = \max\{E_{\text{no issue}}(Y_0, 0), E_{\text{issue}}(Y_0, 0)\}. \tag{3.44}$$

It is well known (see, e.g., DeMarzo and He (2021, Propositions 4 and 6)) that in this case the optimal default boundary is

$$y_{b,\text{Leland}} \equiv \left(\frac{\Pi}{\Pi - 1}\right)\left(\frac{r-\mu}{r+\xi}\right)\left(\frac{c(1-\tau)+\xi}{1-\tau}\right), \tag{3.45}$$

and that the values of debt and equity per unit face value of debt F_t are:

$$p_{\text{Leland}}(y) \;=\; \begin{cases} 0 & \text{if} \quad 0 \leq y \leq y_{b,\text{Leland}} \\ \left(\frac{c+\xi}{r+\xi}\right)\left(1 - \left(\frac{y}{y_{b,\text{Leland}}}\right)^{\Pi}\right) & \text{if} \quad y_t > y_{b,\text{Leland}} \end{cases}, \tag{3.46}$$

$$e_{\text{Leland}}(y) \;=\; \begin{cases} 0 & \text{if} \quad 0 \leq y \leq y_{b,\text{Leland}} \\ \hat{e}(y) - \frac{1-\tau}{(r-\mu)\Pi}\left(\frac{y}{y_{b,\text{Leland}}}\right)^{\Pi} y_{b,\text{Leland}} & \text{if} \quad y_t > y_{b,\text{Leland}} \end{cases}. \tag{3.47}$$

Here, $\hat{e}(y)$ is the value of the levered claim to EBIT as defined in equation (3.28), and the exponent Π is a negative constant defined in equation (3.15).

Substituting the equity and debt values into equation (3.43), we find

$$\left(\frac{\hat{y}_{\text{Leland}}}{y_{b,\text{Leland}}}\right)^{\Pi} \;=\; \frac{\tau c}{\tau c + (-\Pi)(c+\xi)}. \tag{3.48}$$

Finally, setting $E_{\text{no issue}}(Y, 0) = E_{\text{issue}}(Y, 0)$, we find the value of β for which the firm is indifferent to issuing debt or not:

$$\beta^*(\xi) \;=\; \left(\frac{1-\tau}{r-\mu}\right)\left(\frac{\tau c}{\tau c - \Pi(c+\xi)}\right)^{\frac{\Pi-1}{\Pi}} \left[\frac{(c+\xi)(1-\Pi)}{c(1-\tau)+\xi} - 1\right]. \tag{3.49}$$

An important special case is when $\xi = 0$, which determines an upper bound for β above which the optimal upper boundary is infinity regardless of maturity:

$$\overline{\beta} \;\; = \;\; \left(\frac{\tau}{r-\mu}\right)\left(\frac{\tau - \Pi^*}{\tau}\right)^{\frac{1}{\Pi^*}}, \tag{3.50}$$

where the exponent $\Pi^* \equiv \Pi_{(\xi=0)} = \frac{\frac{\sigma^2}{2}-\mu-\sqrt{\left(\frac{\sigma^2}{2}-\mu\right)^2+2r\sigma^2}}{\sigma^2} \; < \; 0.$

As we discuss below in Section 3.3.1, the Leland (1994) benchmark is important because the function $\beta^*(\xi)$ derived in equation (3.49) provides an analytic expression that separates the region of existence of equilibria into MPEs for which the optimal issuance boundary y_u is finite from those for which it is infinite. Moreover, equation (3.48) identifies the limiting value of the optimal target $\hat{y}$ in our model as $y_u \to \infty$.

Generalized Goldstein, Ju, and Leland model

The second benchmark with commitment is a version of the Goldstein, Ju, and Leland (2001) model in which shareholders can commit to an optimal restructuring threshold y_u and target $\hat{y}$, but are still subject to limited liability. The reason we allow for commitment at the upper boundary but not at the lower boundary is because debt issuance is potentially a repeated game in which reputation effects might allow such a policy to be supported within an subgame perfect equilibrium. In contrast, there is no repeated game associated with default.

Proposition 5 (Goldstein, Ju, and Leland model)**.** *Let the income-to-debt ratio dynamics* (dy_t) *be given by the process* (3.6)*, and assume that shareholders can commit to a future debt issuance threshold* y_u *and rebalancing target* $\hat{y}$*. Then the optimal commitment policy* $\mathbf{a}_{GJL} = (y_b, y_u, \hat{y})$ *is given by the following smooth pasting and first-order conditions:*

$$e'(y_b|\mathbf{a}_{GJL}) = 0, \qquad \frac{\partial e(y|\mathbf{a}_{GJL})}{\partial y_u} = 0, \qquad \frac{\partial e(y|\mathbf{a}_{GJL})}{\partial \hat{y}} = 0, \tag{3.51}$$

for any $y \in (y_b, y_u)$*. The equity and debt values are*

$$e(y|\mathbf{a}_{GJL}) \;\; = \;\; \hat{e}(y) + M_e(\mathbf{a}_{GJL})y^{\Theta} + N_e(\mathbf{a}_{GJL})y^{\Pi} \tag{3.52}$$

$$p(y|\mathbf{a}_{GJL}) \;\; = \;\; \frac{c+\xi}{r+\xi} + M_p(\mathbf{a}_{GJL})y^{\Theta} + N_p(\mathbf{a}_{GJL})y^{\Pi}, \tag{3.53}$$

where $\hat{e}(y)$ *is the value of the levered claim to EBIT defined in equation* (3.28)*;* $\Theta > 1$ *and* $\Pi < 0$ *are constants given in equations* (3.14) *and* (3.15)*, and the constants* $M_e(\mathbf{a}_{GJL})$*,* $N_e(\mathbf{a}_{GJL})$*,*

$M_p(\mathbf{a}_{GJL})$, and $N_p(\mathbf{a}_{GJL})$ are uniquely determined by the boundary conditions

$$\begin{aligned}
p(y_b|\mathbf{a}_{GJL}) &= 0 \quad \textit{(zero debt recovery at default)} &(3.54)\\
p(y_u|\mathbf{a}_{GJL}) &= p(\hat{y}|\mathbf{a}_{GJL}) \quad \textit{(no jump at issuance)} &(3.55)\\
e(y_b|\mathbf{a}_{GJL}) &= 0 \quad \textit{(zero equity recovery at default)} &(3.56)\\
e(y_u|\mathbf{a}_{GJL}) &= \frac{y_u}{\hat{y}} e(\hat{y}|\mathbf{a}_{GJL}) + \left(\frac{y_u}{\hat{y}} - 1\right) p(\hat{y}|\mathbf{a}_{GJL}) - \beta y_u \quad \textit{(equity value at issuance)} &(3.57)
\end{aligned}$$

In the Goldstein, Ju, and Leland (2001) model, commitment reflects the shareholders' ability to choose an optimal restructuring policy for the entire life of the firm. This is in contrast to a model without commitment that requires incentive compatibility, that is, shareholders never choose to deviate from the restructuring strategy that creditors use to price the bonds at issuance. A special case of Proposition 5 is the limiting case of vanishing issuance costs:

Corollary 1 (Vanishing issuance costs)**.** *When $\beta \to 0$, the equity and debt values $e(y|\mathbf{a}_{GJL})$ and $p(y|\mathbf{a}_{GJL})$ are given by equations* (3.52)–(3.53) *subject to the default boundary conditions* (3.54) *and* (3.56) *and the restructuring boundary conditions*

$$\begin{aligned}
p'(y_u) &= 0, &(3.58)\\
y_u e'(y_u) &= e(y_u) + p(y_u). &(3.59)
\end{aligned}$$

The boundary condition at equation (3.59) is obtained from the boundary condition (3.57) by setting $\beta = 0$ and performing a first-order Taylor series expansion around the point $(y_u/\hat{y}) - 1$. Because $\beta = 0$, this model has the same setup as that of DeMarzo and He (2021). The next proposition establishes the global optimality of the $\mathbf{a}_{\text{GJL}}$ policy in that the values of $(y_u, \hat{y})$ that maximize the equity claim are independent of the state variable y_t. As such, the $\mathbf{a}_{\text{GJL}}$ commitment policy serves as a natural benchmark for investigating the disparity in tax benefits between models with and without commitment. As we show in Section 3.3.5, the global optimality of the $\mathbf{a}_{\text{GJL}}$ commitment policy implies that such a policy can be a subgame perfect equilibrium when the punishment from deviation is that all future debt issuances will be priced according to the model of DeMarzo and He (2021).

Proposition 6 (Global optimality of the Goldstein, Ju, and Leland policy)**.** *The commitment policy $\mathbf{a}_{GJL}$ defined in Proposition 5 is globally optimal, that is $e(y|\mathbf{a}_{GJL}) = \sup_{\mathbf{a}\in\mathcal{B}} e(y|\mathbf{a})$ for all $y \in (y_b, y_u)$.*

The intuition for this global optimal property can be gleaned from combining the equity valuation given in equation (3.52) with the boundary condition at y_b given in equation (3.56):

$$e(y|\mathbf{a}_{\text{GJL}}) = \left[\hat{e}(y) - \hat{e}(y_b)\left(\frac{y}{y_b}\right)^{\Pi}\right] + M_e(\mathbf{a}_{\text{GJL}})\left[y^{\Theta} - y_b^{\Theta}\left(\frac{y}{y_b}\right)^{\Pi}\right]. \quad (3.60)$$

Importantly, the only place the restructuring parameters $(y_u, \hat{y})$ appear in equation (3.60) is through the coefficient $M_e(\mathbf{a}_{\mathrm{GJL}})$. The implication is that the first order conditions for $(y_u, \hat{y})$ are independent of the state variable y:

$$\begin{aligned} 0 &= \frac{\partial e(y|\mathbf{a}_{\mathrm{GJL}})}{\partial y_u} = \frac{\partial M_e(\mathbf{a}_{\mathrm{GJL}})}{\partial y_u} \\ 0 &= \frac{\partial e(y|\mathbf{a}_{\mathrm{GJL}})}{\partial \hat{y}} = \frac{\partial M_e(\mathbf{a}_{\mathrm{GJL}})}{\partial \hat{y}}. \end{aligned} \tag{3.61}$$

The global-optimality result from Proposition 6 explains why the first-order conditions in equation (3.51) apply for any value $y \in (y_b, y_u)$.

3.3 Results

In this section, we investigate the properties of the MPE in barrier strategies derived from our model. Table 3.1 reports the coefficients for the baseline calibration of the model. We set the annual risk-free rate $r = 4\%$, and the drift and volatility of the EBIT dynamics in equation (3.1) to $\mu = 0$ and $\sigma = 22\%$, respectively. We assume that corporate profits are taxed at a rate $\tau = 20\%$ and we fix the coupon rate $c = r$.[7]

The key parameters of our model are the debt issuance cost parameter β and the inverse maturity parameter ξ. We calibrate them to match two empirical facts: (i) an average debt maturity of three to seven years (e.g., Choi, Hackbarth, and Zechner (2018)) and (ii) debt issuance costs in the range of one to two percent of the amount issued (e.g., Altınkılıç and Hansen (2000)). Within these ranges, we focus on a maturity of five years (i.e., $\xi = 0.2$), and a debt issuance cost parameter β that generates a 1% fee on the amount raised. That is, we choose β so that:

$$\beta\, y_u \overset{\text{set}}{=} \text{target 1\% cost } \times \underbrace{\left(\frac{y_u}{\hat{y}} - 1\right) p(\hat{y})}_{\text{Funds raised}}, \tag{3.62}$$

where $p(\hat{y})$, which equals $p(y_u)$, is the MPE debt issuance price when debt maturity is five years. We obtain a value $\beta = 0.065$, which corresponds to 0.026% of asset value.

3.3.1 Existence of MPEs in barrier strategies

Figure 3.1 partitions the parameter space (β, ξ) into two regions separated by the blue line labeled "existence threshold." The points to the right of this blue line represent MPEs in that, for these points, the necessary and sufficient conditions of Proposition 4 are satisfied. Within this region, the red line labeled "no-issuance threshold" further partitions barrier-strategy MPEs

[7] Choosing c so that the bond is priced at par at issuance generates nearly identical results.

into two sub-regions: one region with a *finite* restructuring boundary y_u, and one region with $y_u = \infty$, implying that firms in this region optimally choose to not issue debt in the future. This red line $\beta^*(\xi)$ is determined analytically by equation (3.49). Figure 3.2 shows that, as the issuance cost parameter β approaches $\beta^*(\xi)$, $(y_b, \hat{y})$ converge to those of the generalized Leland (1994) model, equations (3.45) and (3.48), and y_u approaches infinity. That is, when β is sufficiently high so that the optimal y_u goes to infinity, both the no-commitment model and the global optimal model with commitment converge to the Leland model.

The intuition for the three regions in Figure 3.1 is the following: The roll-over cost of debt for a given maturity $(1/\xi)$ is increasing in the issuance cost parameter β. Hence, for sufficiently large values of β, issuance costs exceed the tax-benefits from debt and, in equilibrium, the firm finds it optimal to not issue debt in the future (the region to the right of the red line in Figure 3.1). A special case in this region is $\overline{\beta} = \lim_{\xi \to 0} \beta^*(\xi)$, defined in equation (3.50) and denoted by the red dotted line, where the optimal upper restructuring boundary y_u is infinity regardless of maturity. For intermediate levels of β, the tax benefit from debt exceeds issuance costs and the MPE is characterized by a barrier strategy with a finite restructuring boundary y_u (the region in between the blue and the red lines).

However, for sufficiently low values of β, there is no barrier strategy MPE (the region to the left of the blue line in Figure 3.1), in spite of the fact that, for optimal policies with commitment, lower values of β offer even larger tax benefits to debt. The lack of MPE in this region is thus due to the myopic nature of no-commitment equilibria, the intuition for which can be gleaned from Figure 3.2. As Panels B and C demonstrate, as the issuance cost parameter declines, management chooses to issue debt more aggressively, both by reducing the debt issuance boundary y_u and the post-issuance target $\hat{y}$. As shown in Panel D, this more aggressive debt issuance policy causes bondholders to reduce the price they are willing to pay for the debt issuance. For sufficiently low values of β, Panel A shows that this lower debt price leads management to default at higher thresholds y_b. As β is lowered from just above the MPE existence threshold to just below it, the reduction in debt issuance price and increase in the location of the default boundary lead to a vicious circle, leading to a situation with no Markov Perfect equilibrium. The location of the MPE existence threshold coincides with the left-most point of the red lines in Figure 3.2

Mathematically, the reason there is a region of no MPE is because the necessary and sufficient conditions of Proposition 4 for the existence of an MPE in barrier strategies are violated. Specifically, we find that necessary conditions in equations (3.36)–(3.37) do not hold. As shown in the appendix, these conditions are equivalent to the supermartingale condition derived in equation (3.39).[8]

[8] Hugonnier, Malamud, and Morellec (2015) find a similar no equilibrium region for the case of

In contrast, if the manager can commit to the global optimal policy of Section 3.2.4, the firm issues debt less aggressively and, as β decreases, the value of debt remains largely unchanged (Figure 3.2, bottom panel). Moreover, in the presence of commitment, there is always an optimal barrier strategy for any level of β. For sufficiently large values of β, the global-optimal barrier policy entails never restructuring existing debt, i.e., $y_u \to \infty$. As stated in Section 3.2.4, the no-debt-issuance threshold for the case with commitment coincides with that of the generalized Leland (1994)model.

3.3.2 Off-equilibrium restructuring policy

For the base case parameters ($\beta = 0.065$, $\xi = 0.2$), Figure 3.3 shows the restructuring function $\mathcal{Y}(y) : [y_u, \infty) \to (y_b, y_u)$, defined in equation (A42). In the MPE, the firm issues debt at $y = y_u$, bringing the income-to-debt ratio to the value $\hat{y} \in (y_b, y_u)$, denoted by the black dot in the figure. Off-equilibrium, i.e., $y > y_u$, the target income-to-debt ratio $\mathcal{Y}(y)$ is an increasing function of y. The dependence of the restructuring target $\mathcal{Y}(y) \in (y_b, y_u)$ on y is a key feature of our model, and is due to our assumption that the firm is not required to repurchase (i.e., call) all outstanding debt prior to issuing new debt. This assumption contrasts with much of the existing literature, which assumes all debt must be called prior to any new debt issuance, which in turn implies that the debt issuance decision is always made at ($y = \infty$), and thus the function $\mathcal{Y}(y)$ reduces to a single value, $\mathcal{Y}(\infty)$.

3.3.3 Tax benefits to debt, issuance costs, and debt maturity

In this section we quantify the tax benefits to debt as a function of restructuring costs and debt maturity (β, ξ). In order to make a fair comparison across MPEs associated with different policy parameters $(y_b, y_u, \hat{y})$, we evaluate the tax benefits for an initially unlevered firm, $F_0 = 0$ (i.e., $y = \infty$). Shareholders choose the initial amount of debt $\hat{F}$ such that:

$$E(Y_0, 0) = \sup_{\hat{F} \geq 0} \left(E(Y_0, \hat{F}) + P(Y_0, \hat{F}) - \beta Y_0 \right). \tag{3.63}$$

We define the tax benefit in terms of: i) the value of the equity claim $E(Y_0, 0)$ for a firm that follows an optimal capital structure policy, and ii) the value of the equity claim for a firm that never issues any debt $E_{\text{no issue}}(Y_0, 0) = (1 - \tau) \left(\frac{Y_0}{r - \mu} \right)$:

$$TB = \left(\frac{E(Y_0, 0) - E_{\text{no issue}}(Y_0, 0)}{E_{\text{no issue}}(Y_0, 0)} \right). \tag{3.64}$$

callable make-whole debt with proportional issuance costs.

As demonstrated in DeMarzo and He (2021), equation (3.64) also applies to the special case $\beta = 0$, in which the tax benefit is zero, regardless of debt maturity.[9]

The left panel of Figure 3.4 shows the tax benefit is zero at the boundary of the MPE existence region, but otherwise is strictly positive with magnitude that varies with debt maturity. In particular, the debt maturity that maximizes tax benefits increases with issuance costs, as highlighted by the red markers. This is due to a tradeoff between debt-issuance fees and tax benefits net of bankruptcy costs. A shorter maturity makes debt less risky and thus allows the manager to increase leverage and extract more tax-benefit. This is evident from equation (3.6) in which the drift of the EBIT process y increases with ξ. Hence, for $\xi \to \infty$ debt becomes virtually risk-free. However, a shorter maturity also implies more frequent rollovers, leading to higher restructuring costs.

To provide a benchmark for the MPE, we also compute the tax benefit in equation (3.64) for the case of the global optimal policy with commitment (blue markers in Figure 3.4). As the figure shows, the value of tax benefit associated with the MPE (red markers) is very close to that of the global-optimal policy (blue markers).

Finally, the right panel of Figure 3.4 shows the equity share of the claim to EBIT, $E(Y_0, 0)/V_0$. Along the (β, ξ) points that maximizes the tax benefit in the MPE, the firm extracts most of the value of the EBIT claim. Indeed, as β is lowered, for both the case with commitment (blue markers) and without commitment (red markers), the equity share $E(Y_0, 0)/V_0$ approaches 1, which, from equation (3.64), corresponds to a tax benefit equal to $\tau/(1-\tau) = 0.25$ in our calibration. In sum, the presence of even arbitrarily small issuance costs can break the irrelevance of capital structure and maturity choices found in DeMarzo and He (2021) that arise in an MPE without issuance costs.

3.3.4 Debt and equity values

The left panel of Figure 3.5 shows the value of equity and debt in the MPE as a function of the income to debt ratio y. As in the baseline calibration, we set $\beta = 0.065$ and $\xi = 0.2$, which implies a fractional cost of 1% of the debt amount issued and a maturity of five years. As shown in Proposition 10, the value of equity associated with the MPE (red line) is nonnegative, nondecreasing, and convex. The black line shows that the bond price in the MPE is strictly concave in the inaction region. For large values of y, the value of debt approaches the price of a risk-free bond with the same promised cash flows, which is $\frac{c+\xi}{r+\xi} = 1$ in our calibration. As y approaches y_b, the debt price drops to zero because there is no recovery at default.

The right panel shows that the difference between equity value in the global optimal policy

[9] See equation (31) of DeMarzo and He (2021), evaluated in the limit $y \to \infty$.

and in the MPE is within 0.3% of the asset value. Related, we find that dynamic debt issuance increases enterprise value by 5.2% compared to an unlevered firm. In the global optimal benchmark with commitment the tax benefit is very close at 6%. This suggests that, when the rebalancing cost β and the inverse maturity ξ are calibrated to match empirical observation, shareholders' inability to commit to a future issuance policy has limited impact on the tax benefits that a firm can extract.

3.3.5 Non-Markov subgame perfect equilibria

So far we have used the global optimal policy with commitment only as a benchmark to study the properties of barrier-strategy MPEs. Here we discuss conditions under which the global optimal policy is a subgame-perfect equilibrium, albeit outside the Markov class. As an example, we consider the case of vanishing issuance cost β in Corollary 1, which allows us to compare the global optimal policy with the DeMarzo and He (2021) smooth issuance policy in which $\beta = 0$. For both cases we consider the empirically relevant five-year debt maturity, i.e., $\xi = 0.2$.

Figure 3.6 shows the value of the debt, $p(y)$, (left panel) and equity, $e(y)$, (right panel) claims. The blue line reports values associated with the global optimal barrier strategy with commitment defined in Corollary 1. The dashed-red line refers to the DeMarzo-He MPE. The figure shows that, for all values of y, the value of debt and equity are always higher under the global optimal policy. Since equityholders are always better off by following the global optimal policy, they do not have incentive to ever deviate if the punishment is that, following a deviation, debt is always priced according to the DeMarzo-He MPE. Because debtholders observe the size of the debt issuance and pay fair value for their claim at the restructuring date, they are indifferent to the firm's debt issuance policy. Therefore, debtholders can credibly threaten to punish any deviation from the optimal policy by pricing debt according to the DeMarzo-He MPE (or the MPE that we identify when β is not zero). As a result, shareholders would gain zero cash benefit at the date of the deviation, and would be left with an equity claim that has lower valuation than under the global optimal policy.[10] Under this "grim-trigger" punishment, the global optimal policy with commitment is a subgame perfect equilibrium outside of the Markov class.[11]

[10] Shareholders would obtain zero cash benefit because debt issuance at the deviation date is a locally deterministic process.

[11] Because debt issuance is a repeated game, managers care about the firm's reputation and target minimum credit rating levels (e.g., Kisgen (2006, 2009)). This provides empirical support for the theoretical argument that optimal policies with commitment can be time-consistent.

3.4 Conclusion

Within a standard tradeoff setting, we derive optimal dynamic capital structure policies when the firm faces fixed debt-restructuring costs, and shareholders cannot commit to future debt policies. We provide necessary and sufficient conditions for the existence of Markov Perfect Equilibria in barrier strategies when the firm is not required to repurchase all of its outstanding debt prior to issuing additional debt. We show that the interaction between issuance costs and debt maturity determines three regions of interest: when issuance costs are sufficiently high and maturity sufficiently low, the MPE is characterized by no future debt issuances. For lower issuance costs and longer maturities, it is optimal for the firm to issue additional debt when leverage becomes sufficiently low, which allows shareholders to extract positive tax benefits. Finally, for even lower issuance costs and longer maturities, we identify a region in which no MPE exists in spite of significant tax benefits that are available to a model with commitment. The lack of an MPE in this region is due to the myopic nature of no-commitment strategies.

We also investigate a barrier model with commitment. We show that the optimal policy is in fact *globally optimal* in that equity values are higher in every state of nature compared to any other model consistent with limited liability. Moreover, we show that this policy can be supported within a subgame perfect equilibria if creditors (credibly) threaten to follow a grim trigger policy in which any deviation leads to all future debt issuances being priced according to the MPE policy. This global optimal equilibrium provides a useful benchmark for estimating how much shareholders are hurt by not being able to commit to a debt issuance policy. Calibrating our model to empirically relevant values (e.g., bonds issued with a five-year maturity, and subject to a 1% debt issuance cost), we find that tax benefits to debt for the no-commitment case are very close to those for the global optimal solution. Finally, as the issuance cost parameter approaches zero, shareholder can extract 100% of the firms EBIT flow under both models with and without commitment. Thus, even for a case of vanishing, but positive, issuance costs, we conclude that not being able to commit to a dynamic capital structure policy may have minimal impact on a firm's ability to extract tax benefits to debt. This is in contrast to the case of zero issuance costs, in which, as shown by DeMarzo and He (2021), there are no tax benefits to debt regardless of maturity choice.

Table 3.1: **Baseline model coefficients**.

The table shows the values of the model coefficients in the baseline calibration. An inverse debt maturity $\xi = 0.2$ corresponds to a five-year expected maturity. The fixed issuance cost $\beta = 0.065$ corresponds to a 1% of the debt amount issued.

Parameter	Symbol	Value
Annual risk-free rate	r	0.04
Annual coupon rate	c	0.04
Annual EBIT drift	μ	0
Annual EBIT volatility	σ	0.22
Corporate tax rate	τ	0.2
Loss given default	α	1
Inverse debt maturity	ξ	0.2
Fixed issuance cost	β	0.065

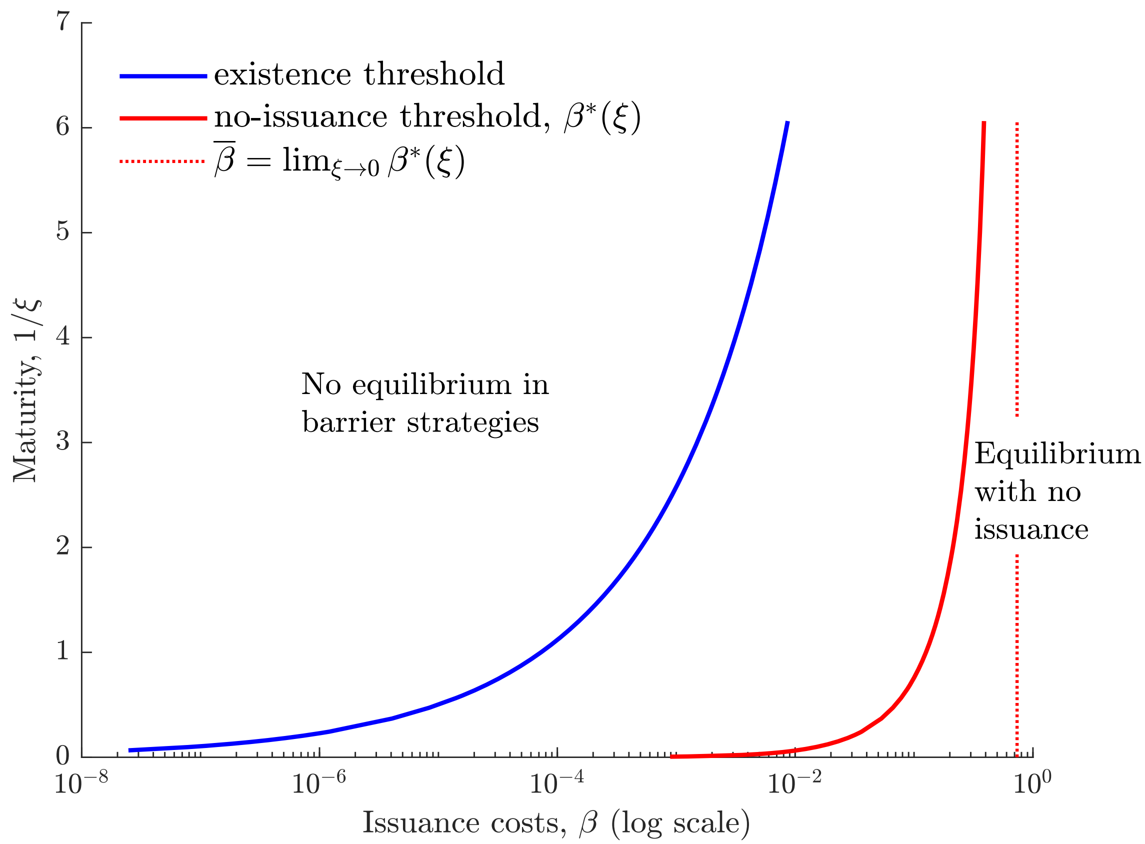


Figure 3.1: Existence of barrier-strategy MPEs.

The figure identifies the regions in the space (β, ξ) of issuance costs and inverse-maturity parameters for which barrier-strategy MPEs exist. In the region to the left of the blue line there is no MPE in barrier strategies. To the right of the red line it is optimal not to issue debt.

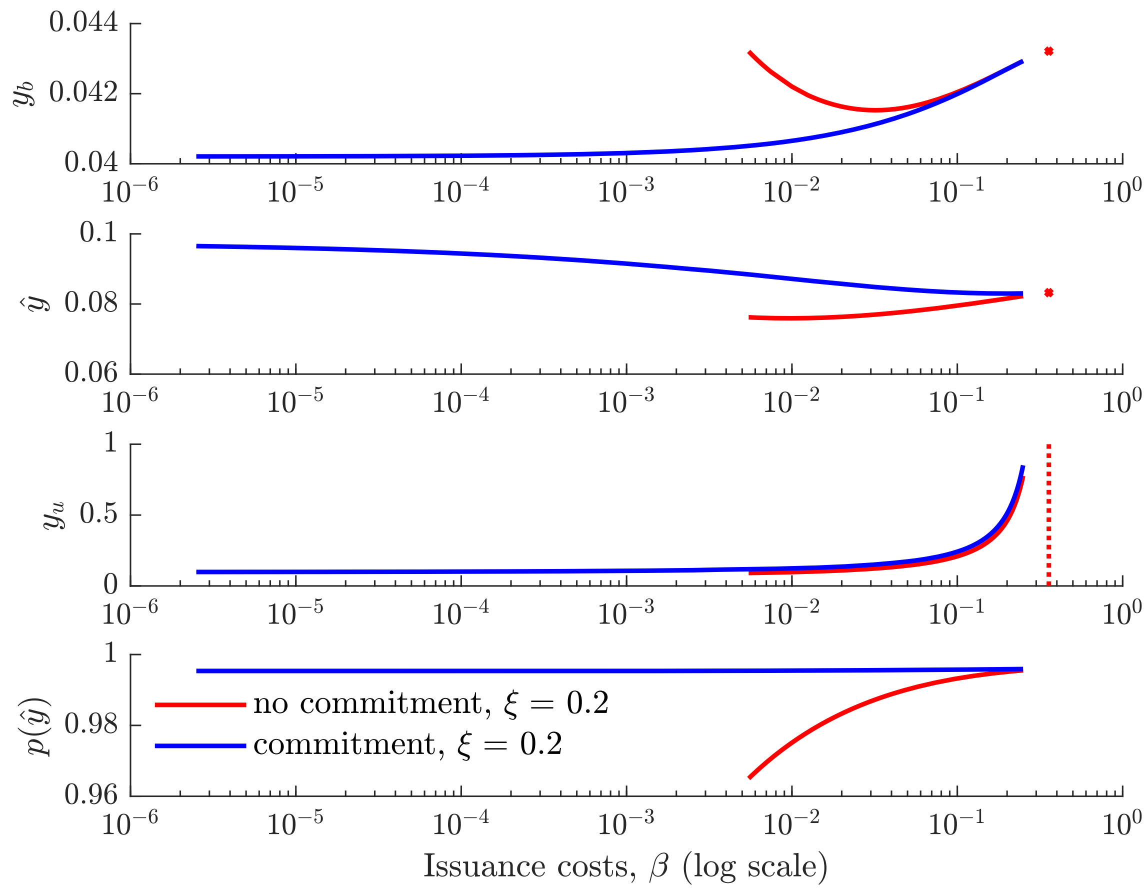


Figure 3.2: Debt issuance policies and values for a fixed maturity.

The figure shows the default boundary y_b (first panel); the post-issuance target $\hat{y}$ (second panel); the restructuring boundary y_u (third panel); and the value of debt $p(\hat{y})$ (fourth panel) as a function of the issuance cost β. The red lines correspond to the no commitment case (MPE) while the blue lines refer to the case of commitment. The two red markers in the top two panels denote the values of y_b and $\hat{y}$ in the Leland model. The vertical dotted line in the next panel refers to $\beta^*(\xi = 0.2)$, which is the smallest value of β for which the optimal issuance boundary is $y_u = \infty$ when $\xi = 0.2$. The inverse maturity parameter is fixed at $\xi = 0.2$, corresponding to a debt maturity of 5 years.

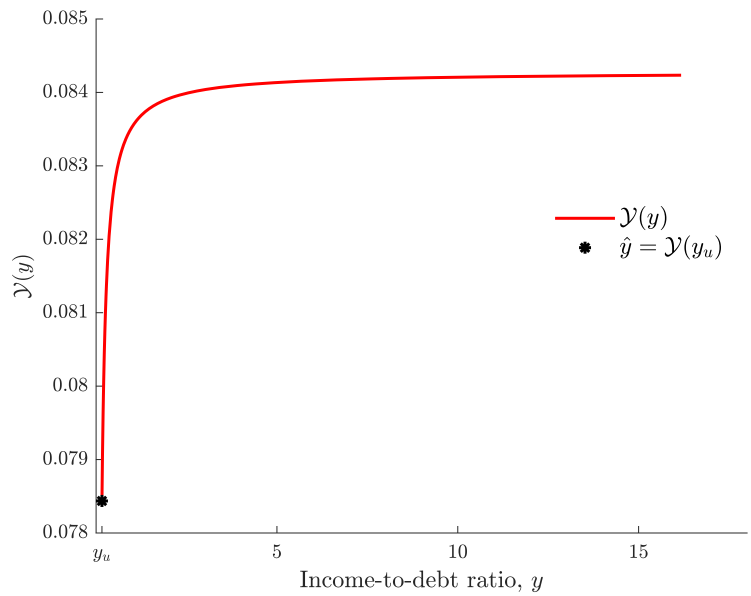


Figure 3.3: Restructuring function $\mathcal{Y}(y)$.

The figure shows the restructuring function $\mathcal{Y}(y)$, $y \geq y_u$, defined in equation (A42). As in the baseline calibration, we set $\beta = 0.065$ and $\xi = 0.2$, which reflect a fractional cost of 1% of the debt amount issued and a maturity of five years. The black dot denotes the restructuring point $(y_u, \hat{y} = \mathcal{Y}(y_u))$.

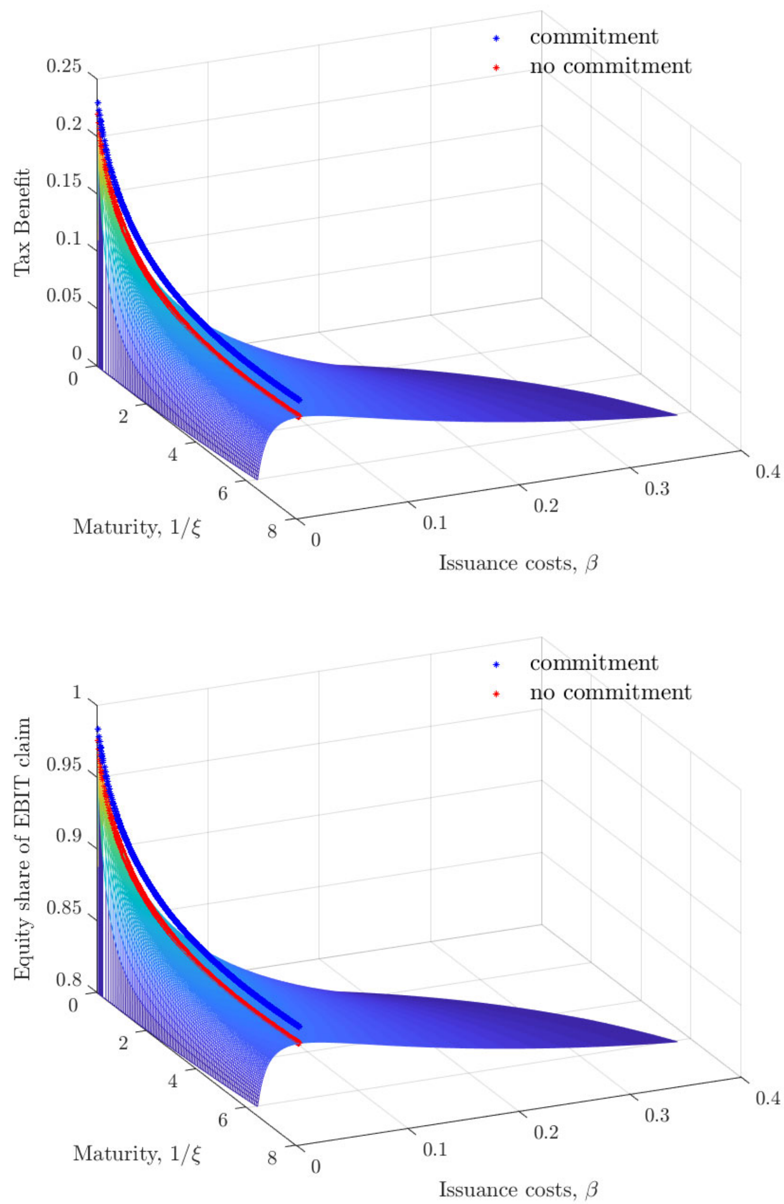


Figure 3.4: Tax benefits and equity share of EBIT claim.

The figure shows the tax benefit (left panel) and the equity share of EBIT (right panel), for values of (β, ξ) that belong to the MPE existence region of Figure 3.1. The red markers correspond to the set (β, ξ) at which the tax benefit in the MPE is highest. For the same set of (β, ξ) points, the blue markers show the tax benefit (left panel) and equity share of EBIT (right panel) under the global optimal policy with commitment.

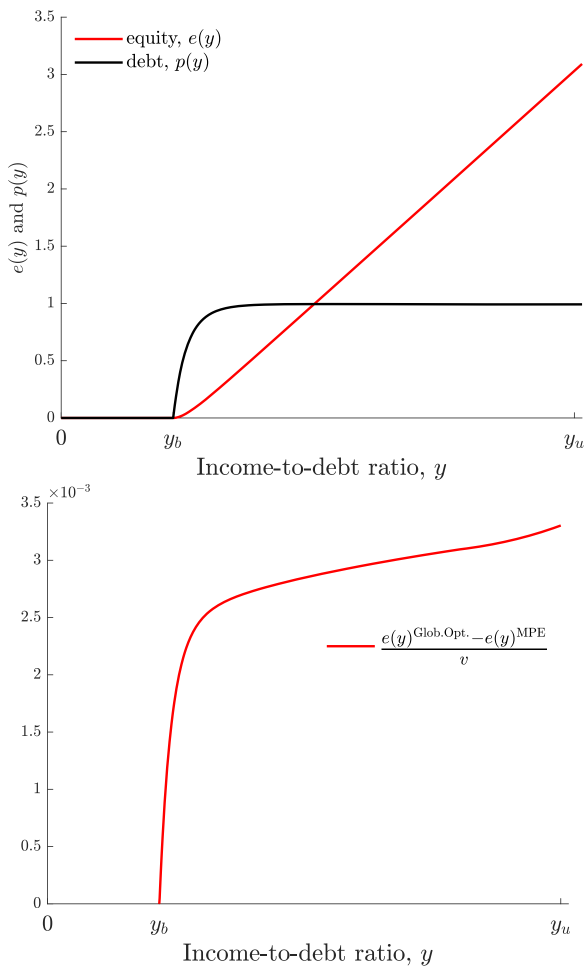


Figure 3.5: Debt and equity values

The left panel shows the value of equity and debt in the MPE as a function of the income to debt ratio y. The right panel shows the difference between equity value in the global optimal policy and in the MPE, scaled by asset value. As in the baseline calibration, we set $\beta = 0.065$ and $\xi = 0.2$, which reflect a fractional cost of 1% of the debt amount issued and a maturity of five years.

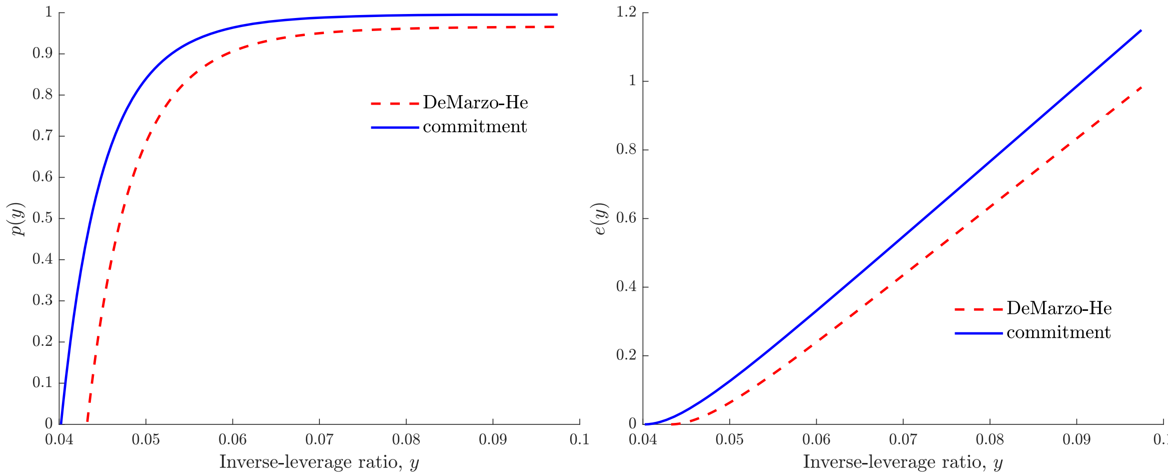


Figure 3.6: Debt and equity values with vanishing issuance costs.

For $\xi = 0.2$, the figure shows the value of the debt $p(y)$ (left panel) and equity $e(y)$ (right panel) claims. The blue line reports values associated with the global optimal barrier strategy with commitment defined in Corollary 1. The dashed-red line refers to the DeMarzo and He (2021) MPE with $\xi = 0.2$.

References

Abdi, F., and B. Wu, 2018, "Informed Corporate Credit Market before Monetary Policy Surprises: Explaining Pre-FOMC Stock Market Movements," *University of St. Gallen, School of Finance Research Paper*, (2018/28).

Abel, A. B., 2016, "Investment with Leverage," Working Paper, Wharton School of the University of Pennsylvania.

Abel, A. B., and J. C. Eberly, 1994, "A Unified Model of Investment Under Uncertainty," *American Economic Review*, 84(1), 1369–1384.

Acharya, V., S. A. Davydenko, and I. A. Strebulaev, 2012, "Cash Holdings and Credit Risk," *Review of Financial Studies*, 25(12), 3572–3609.

Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll, forthcoming, "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *Review of Economic Studies*.

Admati, A. R., P. M. DeMarzo, M. F. Hellwig, and P. Pfleiderer, 2018, "The leverage ratchet effect," *The Journal of Finance*, 73(1), 145–198.

Agapova, A., and J. Madura, 2011, "Information leakage prior to company issued guidance," *Financial Management*, 40(3), 623–646.

Ahern, K. R., 2020, "Do proxies for informed trading measure informed trading? Evidence from illegal insider trades," *The Review of Asset Pricing Studies*, 10(3), 397–440.

Ai, H., and R. Bansal, 2018, "Risk preferences and the macroeconomic announcement premium," *Econometrica*, 86(4), 1383–1430.

Ai, H., R. Bansal, J. Im, and C. Ying, 2018, "A Model of the Macroeconomic Announcement Premium with Production," .

Ai, H., and R. Li, 2015, "Investment and CEO Compensation under Limited Commitment," *Journal of Financial Economics*, 116(3), 452–472.

Albuquerque, R., and H. Hopenhayn, 2004, "Optimal Lending Contracts and Firm Dynamics," *Review of Economic Studies*, 71, 285–315.

Altinkiliç, O., and R. S. Hansen, 2000, "Are There Economies of Scale in Underwriting Fees?," *Review of Financial Studies*, 13(1), 191–218.

Altınkılıç, O., and R. S. Hansen, 2000, "Are there economies of scale in underwriting fees? Evidence of rising external financing costs," *Review of Financial Studies*, 13(1), 191–218.

Alvarez, L. H., et al., 2003, "On the properties of r-excessive mappings for a class of diffusions," *The annals of applied probability*, 13(4), 1517–1533.

Asker, J., A. Collard-Wexler, and J. de Loecker, 2014, "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 122(5), 1013–1063.

Augustin, P., M. Brenner, and M. G. Subrahmanyam, 2015, "Informed options trading prior to M&A announcements: Insider trading?," *Available at SSRN 2441606*.

Back, K., 1992, "Insider trading in continuous time," *The Review of Financial Studies*, 5(3), 387–409.

Back, K., C. H. Cao, and G. A. Willard, 2000, "Imperfect competition among informed traders," *The journal of finance*, 55(5), 2117–2155.

Back, K., K. Crotty, and T. Li, 2018, "Identifying information asymmetry in securities markets," *The Review of Financial Studies*, 31(6), 2277–2325.

Back, K., and H. Pedersen, 1998, "Long-lived information and intraday patterns," *Journal of financial markets*, 1(3-4), 385–402.

Baker, M., and J. Wurgler, 2002, "Market timing and capital structure," *The Journal of Finance*, 57(1), 1–32.

Barclay, M. J., and C. W. Smith, 1995, "The maturity structure of corporate debt," *Journal of Finance*, 50.2, 609–631.

Baruch, S., 2002, "Insider trading and risk aversion," *Journal of Financial Markets*, 5(4), 451–464.

Bates, T. W., K. M. Kahle, and R. M. Stulz, 2009, "Why Do U.S. Firms Hold So Much More Cash Than They Used To?," *Journal of Finance*, 64(5), 1985–2021.

Belo, F., X. Lin, and F. Yang, 2018, "External Equity Financing Shocks, Financial Flows, and Asset Prices," *Review of Financial Studies*, forthcoming.

Benzoni, L., L. Garlappi, and R. S. Goldstein, 2020, "Incomplete Information, Dynamic Debt Issuance, and the Term Structure of Credit Spreads," working paper, Working paper. The Federal Reserve Bank of Chicago.

Bernile, G., J. Hu, and Y. Tang, 2016, "Can information be locked up? Informed trading ahead of macro-news announcements," *Journal of Financial Economics*, 121(3), 496–520.

Biais, B., T. Mariotti, G. Plantin, and J.-C. Rochet, 2007, "Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications," *Review of Economic Studies*, 74, 345–390.

Bils, M., P. J. Klenow, and C. Ruane, 2020, "Misallocation or Mismeasurement?," NBER Working Paper No. 26711.

Bilyi, M., 2018, "Pre-FOMC announcement drift and speculative trading," .

Blouin, J., J. E. Core, and W. Guay, 2010, "Have the tax benefits of debt been overestimated?," *Journal of Financial Economics*, 98(2), 195–213.

Boguth, O., V. Gregoire, and C. Martineau, 2019, "Shaping expectations and coordinating attention: The unintended consequences of FOMC press conferences," *Journal of Financial and Quantitative Analysis*, 54(6), 2327–2353.

Bolton, P., H. Chen, and N. Wang, 2011a, "A Unified Theory of Tobin's q, Corporate Investment, Financing, and Risk Management," *Journal of Finance*, 66(5), 1545–1578.

——— , 2011b, "A unified theory of Tobin's q, corporate investment, financing, and risk management," *The journal of Finance*, 66(5), 1545–1578.

——— , 2013, "Market Timing, Investment, and Risk Management," *Journal of Financial Economics*, 109, 40–62.

Bolton, P., N. Wang, and J. Yang, 2020, "Leverage Dynamics and Financial Flexibility," working paper, National Bureau of Economic Research.

Boualam, Y., 2019, "Credit Markets and Relationship Capital," Working paper.

Boualam, Y., J. F. Gomes, and C. Ward, 2020, "Understanding the Behavior of Distressed Stocks," Working paper.

Brunnermeier, M. K., and M. Oehmke, 2013, "The maturity rat race," *The Journal of Finance*, 68(2), 483–521.

Brunnermeier, M. K., and M. Yogo, 2009, "A note on liquidity risk management," *American Economic Review*, 99(2), 578–83.

Buera, F. J., J. P. Kaboski, and Y. Shin, 2011, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, 101(5), 1964–2002.

Caldentey, R., and E. Stacchetti, 2010, "Insider trading with a random deadline," *Econometrica*, 78(1), 245–283.

Calomiris, C. W., and M. Tsoutsoura, 2013, "Underwriting Costs of Seasoned Equity Offerings: Cross-Sectional Determinants, Technological Change, and Pricing Benefits, 1980–2008," Working Paper.

Campbell, J. Y., C. E. Pflueger, and L. M. Viceira, 2020, "Macroeconomic drivers of bond and equity risks," .

Campbell, J. Y., A. Sunderam, and L. M. Viceira, 2017, "Inflation bets or deflation hedges? The changing risks of nominal bonds," working paper.

Chen, H., 2010, "Macroeconomic Conditions and the Puzzles of Credit Spreads and Capital Structure," *The Journal of Finance*, 65(6), 2171–2212.

Chen, H., Y. Xu, and J. Yang, 2012, "Systematic risk, debt maturity, and the term structure of credit spreads," working paper, National Bureau of Economic Research.

Cheng, I.-H., and K. Milbradt, 2012, "The hazards of debt: Rollover freezes, incentives, and bailouts," *The Review of Financial Studies*, 25(4), 1070–1110.

Choi, J., D. Hackbarth, and J. Zechner, 2018, "Corporate debt maturity profiles," *Journal of Financial Economics*, 130(3), 484–502.

Cieslak, A., A. Morse, and A. Vissing-Jorgensen, 2019, "Stock returns over the FOMC cycle," *The Journal of Finance*, 74(5), 2201–2248.

Cieslak, A., and H. Pang, 2020, "Common shocks in stocks and bonds," .

Clementi, G. L., and H. A. Hopenhayn, 2006, "A Theory of Financing Constraints and Firm Dynamics," *Quarterly Journal of Economics*, 121, 229–265.

Cochrane, J. H., 2017, "Macro-finance," *Review of Finance*, 21(3), 945–985.

Cocoma, P., 2018, "Explaining the pre-announcement drift," *Available at SSRN 3014299.*

Collin-Dufresne, P., and V. Fos, 2015, "Do prices reveal the presence of informed trading?," *The Journal of Finance*, 70(4), 1555–1582.

——— , 2016, "Insider trading, stochastic liquidity, and equilibrium prices," *Econometrica*, 84(4), 1441–1475.

Crandall, M. G., and P.-L. Lions, 1983, "Viscosity solutions of Hamilton-Jacobi equations," *Transactions of the American mathematical society*, 277(1), 1–42.

Crego, J. A., 2020, "Why does public news augment information asymmetries?," *Journal of Financial Economics.*

Dai, L., Y. Wang, and M. Yang, 2019, "Insider trading when there may not be an insider," *Available at SSRN 2720736.*

Dangl, T., and J. Zechner, 2020, "Debt maturity and the dynamics of leverage," *Review of Financial Studies, Forthcoming.*

Danis, A., D. A. Rettl, and T. M. Whited, 2014, "Refinancing, profitability, and capital structure," *Journal of Financial Economics*, 114(3), 424–443.

Décamps, J.-P., and S. Villeneuve, 2014, "Rethinking Dynamic Capital Structure Models with Roll-Over Debt," *Mathematical Finance*, 24, 66–96.

Della Seta, M., E. Morellec, and F. Zucchi, 2020, "Short-term debt and incentives for risk-taking," *Journal of Financial Economics.*

DeMarzo, P., and Z. He, 2021, "Leverage Dynamics without Commitment," *Journal of Finance*, 76(3), 1195–1250.

DeMarzo, P., Z. He, and F. Tourre, 2018, "Sovereign Debt Ratchets and Welfare Destruction," Working paper, Stanford University, University of Chicago, and Copenhagen Business School.

DeMarzo, P. M., 2019, "Presidential Address: Collateral and Commitment," *The Journal of Finance*, 74(4), 1587–1619.

DeMarzo, P. M., and M. Fishman, 2007, "Agency and Optimal Investment Dynamics," *Review of Financial Studies*, 20, 151–188.

DeMarzo, P. M., M. J. Fishman, Z. He, and N. Wang, 2012, "Dynamic Agency and the q Theory of Investment," *Journal of Finance*, 67(6), 2295–2340.

DeMarzo, P. M., and Y. Sannikov, 2006, "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model," *Journal of Finance*, 61(6), 2681–2724.

Diamond, D. W., and Z. He, 2014, "A theory of debt maturity: the long and short of debt overhang," *The Journal of Finance*, 69(2), 719–762.

Dittmar, A., J. Mahrt-Smith, and H. Servaes, 2003, "International Corporate Governance and Corporate Cash Holdings," *Journal of Financial and Quantitative Analysis*, 38(1), 111–133.

Donaldson, G., 1984, *Managing Corporate Wealth*. New York: Praeger.

Dow, J., G. Gorton, and A. Krishnamurthy, 2005, "Equilibrium Investment and Asset Prices under Imperfect Corporate Control," *American Economic Review*, 95(3), 659–681.

Drechsler, I., A. Moreira, and A. Savov, 2018, "Liquidity creation as volatility risk," *Available at SSRN 3133291*.

Eberly, J., S. Rebelo, and N. Vincent, 2009, "Investment and Value: A Neoclassical Benchmark," NBER Working Paper No. 13866.

Edmans, A., X. Gabaix, and A. Landier, 2008, "A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium," *Review of Financial Studies*, 22(12), 4881–4917.

Erickson, T., and T. M. Whited, 2000, "Measurement Error and the Relationship between Investment and q," *Journal of Political Economy*, 108(5), 1027–1057.

Ernst, R., T. Gilbert, and C. M. Hrdlicka, 2019, "More than 100% of the equity premium: How much is really earned on macroeconomic announcement days?," *Available at SSRN 3469703*.

Fama, E. F., and K. R. French, 2002, "Testing trade-off and pecking order predictions about dividends and debt," *The Review of Financial Studies*, 15(1), 1–33.

Fama, E. F., and K. R. French, 2005, "Financing Decisions: Who Issues Stock?," *Journal of Financial Economics*, 76, 549–582.

Farre-Mensa, J., R. Michaely, and M. Schmalz, 2014, *Payout Policy*vol. 6, pp. 75–134.

Fazzari, S. M., R. G. Hubbard, and B. C. Petersen, 1988, "Financing Constraints and Corporate Investment," *Brookings Papers On Economic Activity*, (1), 141–206.

Finer, D. A., 2018, "What Insights Do Taxi Rides Offer into Federal Reserve Leakage?," *Chicago Booth: George J. Stigler Center for the Study of the Economy & the State Working Paper*, (18).

Fischer, E. O., R. Heinkel, and J. Zechner, 1989, "Dynamic capital structure choice: Theory and tests," *The Journal of Finance*, 44(1), 19–40.

Foster, F. D., and S. Viswanathan, 1996, "Strategic trading when agents forecast the forecasts of others," *The Journal of Finance*, 51(4), 1437–1478.

Frank, M. Z., and V. K. Goyal, 2014, "The profits–leverage puzzle revisited," *Review of Finance*, 19(4), 1415–1453.

Giacoletti, M., R. Ramcharan, and E. Yu, 2020, "The FOMC Cycle and Consumer Decision Making: Evidence from the Residential Housing Market," *Available at SSRN 3364588.*

Gillan, S. L., J. C. Hartzell, A. Koch, and L. T. Starks, 2018, "Getting the Incentives Right: Backfilling and Biases in Executive Compensation Data," *Review of Financial Studies*, 31(4), 1460–1498.

Glover, B., and O. Levine, 2017, "Idiosyncratic Risk and the Manager," *Journal of Financial Economics*, 126(2), 320–341.

Goldstein, R., N. Ju, and H. Leland, 2001, "An EBIT-Based Model of Dynamic Capital Structure," *The Journal of Business*, 74(4), 483–512.

Gomes, J. F., 2001, "Financing Investment," *American Economic Review*, 91(5), 1263–1285.

Gomes, J. F., and L. Schmid, 2012, "Equilibrium Credit Spreads and the Macroeconomy," Working Paper, Duke University.

Gopinath, G., S. Kalemli-Ozcan, L. Karabarbounis, and C. Villegas-Sanchez, 2017, "Capital Allocation and Productivity in South Europe," *Quarterly Journal of Economics*, 132(4), 1915–1967.

Graham, J. R., and C. R. Harvey, 2001, "The theory and practice of corporate finance: Evidence from the field," *Journal of Financial Economics*, 60(2-3), 187–243.

Hall, R. E., 2001, "Struggling to Understand the Stock Market," *American Economic Review*, 91(2), 1–11.

Haltiwanger, J., R. Kulick, and C. Syverson, 2018, "Misallocation Measures: The Distortion that Ate the Residual," NBER Working Paper No. 24199.

Hansen, L. P., and R. Jagannathan, 1991, "Implications of security market data for models of dynamic economies," *Journal of political economy*, 99(2), 225–262.

Hartman-Glaser, B., S. Mayer, and K. Milbradt, 2019, "Corporate Liquidity Management under Moral Hazard," Working Paper.

Hayashi, F., 1982, "Tobin's Marginal q and Average Q: A Neoclassical Formulation," *Econometrica*, 50, 215–224.

He, Z., and K. Milbradt, 2014, "Endogenous liquidity and defaultable bonds," *Econometrica*, 82(4), 1443–1508.

——— , 2016, "Dynamic debt maturity," *The Review of Financial Studies*, 29(10), 2677–2736.

He, Z., and W. Xiong, 2012a, "Debt financing in asset markets," *American Economic Review*, 102(3), 88–94.

——— , 2012b, "Dynamic debt runs," *The Review of Financial Studies*, 25(6), 1799–1843.

Hennessey, C. A., and T. M. Whited, 2007, "How Costly is External Financing? Evidence from a Structural Estimation," *Journal of Finance*, 62, 1705–1745.

Hennessy, C. A., and T. M. Whited, 2007, "How Costly Is External Financing? Evidence from a Structural Estimation," *The Journal of Finance*, 62(4), 1705–1745.

Ho, T., and H. R. Stoll, 1981, "Optimal dealer pricing under transactions and return uncertainty," *Journal of Financial economics*, 9(1), 47–73.

Hopenhayn, H. A., 1992, "Entry, Exit, and Firm Dynamics in Long-Run Equilibrium," *Econometrica*, 60(5), 1127–1150.

Hsieh, C.-T., and P. J. Klenow, 2009, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124(4), 1403–1448.

Hu, G. X., J. Pan, J. Wang, and H. Zhu, 2019, "Premium for heightened uncertainty: Solving the fomc puzzle," working paper, National Bureau of Economic Research.

Hugonnier, J., S. Malamud, and E. Morellec, 2015, "Credit market frictions and capital structure dynamics," *Journal of Economic Theory*, 157, 1130–1158.

Jarocinski, M., and P. Karadi, 2020, "Deconstructing monetary policy surprises—the role of information shocks," *American Economic Journal: Macroeconomics*, 12(2), 1–43.

Jarrell, G. A., and A. B. Poulsen, 1989, "Stock trading before the announcement of tender offers: Insider trading or market anticipation," *JL Econ. & Org.*, 5, 225.

Jensen, M. C., 1986, "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers," *American Economic Review, Papers and Proceedings*, pp. 323–329.

Jensen, M. C., and W. H. Meckling, 1976, "Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure," *Journal of Financial Economics*, pp. 305–360.

Kane, A., A. J. Marcus, and R. L. McDonald, 1984, "How big is the tax advantage to debt?," *The Journal of Finance*, 39(3), 841–853.

———, 1985, "Debt policy and the rate of return premium to leverage," *Journal of Financial and Quantitative Analysis*, 20(4), 479–499.

Keown, A. J., and J. M. Pinkerton, 1981, "Merger announcements and insider trading activity: An empirical investigation," *The journal of finance*, 36(4), 855–869.

Kermani, A., and Y. Ma, 2020, "Asset Specificity of Non-Financial Firms," Working Paper.

King, R. G., and R. Levine, 1993, "Finance and Growth: Schumpeter Might Be Right," *Quarterly Journal of Economics*, 108(3), 717–737.

Kisgen, D. J., 2006, "Credit ratings and capital structure," *The Journal of Finance*, 61(3), 1035–1072.

———, 2009, "Do firms target credit ratings or leverage levels?," *Journal of Financial and Quantitative Analysis*, 44(6), 1323–1344.

Kruk, L., J. Lehoczky, K. Ramanan, and S. Shreve, 2007, "An explicit formula for the Skorokhod map on [0, a]," *The Annals of Probability*, pp. 1740–1768.

Krusell, P., and A. A. Smith, 1998, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(5), 867–896.

Kurov, A., A. Sancetta, G. Strasser, and M. H. Wolfe, 2019, "Price drift before US macroeconomic news: Private information about public announcements?," *Journal of Financial and Quantitative Analysis*, 54(1), 449–479.

Kurov, A., M. Wolfe, and T. Gilbert, 2020, "The Disappearing Pre-FOMC Announcement Drift," *Available at SSRN 3134546*.

Kyle, A. S., 1985, "Continuous auctions and insider trading," *Econometrica: Journal of the Econometric Society*, pp. 1315–1335.

Laarits, T., 2019, "Pre-Announcement Risk," *Available at SSRN 3443886*.

Lamberton, D., M. Zervos, et al., 2013, "On the optimal stopping of a one-dimensional diffusion," *Electronic Journal of Probability*, 18.

Larcker, D. F., 1983, "The Association Between Performance Plan Adoption and Corporate Capital Investment," *Journal of Accounting and Economics*, 5, 3–30.

Leary, M. T., and M. R. Roberts, 2005, "Do firms rebalance their capital structures?," *The Journal of Finance*, 60(6), 2575–2619.

Leland, H. E., 1994, "Corporate debt value, bond covenants, and optimal capital structure," *The Journal of Finance*, 49(4), 1213–1252.

———, 1998, "Agency costs, risk management, and capital structure," *The Journal of Finance*, 53(4), 1213–1243.

Leland, H. E., and K. B. Toft, 1996, "Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads," *The Journal of Finance*, 51(3), 987–1019.

Li, T., 2013, "Insider trading with uncertain informed trading," *Available at SSRN 946324*.

Lintner, J., 1956, "Distribution of Incomes of Corporations Among Dividends, Retained Earnings, and Taxes," *American Economic Review*, 46(2), 97–113, Papers and Proceedings of the Sixty-Eighth Annual Meeting of the American Economic Association.

Lucca, D. O., and E. Moench, 2015, "The pre-FOMC announcement drift," *The Journal of Finance*, 70(1), 329–371.

Lustig, H., C. Syverson, and S. Van Nieuwerburgh, 2011, "Technological Change and the Growing Inequality in Managerial Compensation," *Journal of Financial Economics*, 99, 601–627.

Malenko, A., and A. Tsoy, 2020, "Optimal Time-Consistent Debt Policies," Working paper, Boston College.

McKeon, S. B., 2015, "Employee Option Exercise and Equity Issuance Motives," Working Paper.

Mehra, R., and E. C. Prescott, 1985, "The equity premium: A puzzle," *Journal of monetary Economics*, 15(2), 145–161.

Meulbroek, L. K., 1992, "An empirical analysis of illegal insider trading," *The Journal of Finance*, 47(5), 1661–1699.

Midrigan, V., and D. Y. Xu, 2014, "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, 104(2), 422–458.

Milbradt, K., and M. Oehmke, 2015, "Maturity Rationing and Collective Short-Termism," *Journal of Financial Economics*, 118(3), 553–570.

Milgrom, P., and I. Segal, 2002, "Envelope theorems for arbitrary choice sets," *Econometrica*, 70(2), 583–601.

Moll, B., 2014, "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?," *American Economic Review*, 104(10), 3186–3221.

Morellec, E., B. Nikolov, and N. Schürhoff, 2012, "Corporate governance and capital structure dynamics," *The Journal of Finance*, 67.3, 803–848.

Morellec, E., B. Nikolov, and N. Schürhoff, forthcoming, "Agency Conflicts Around the World," *Review of Financial Studies.*

Myers, S. C., and N. S. Majluf, 1984, "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have," *Journal of Financial Economics*, 13, 187–221.

Nakamura, E., and J. Steinsson, 2018, "High-frequency identification of monetary non-neutrality: the information effect," *The Quarterly Journal of Economics*, 133(3), 1283–1330.

Nikolov, B., and T. M. Whited, 2014, "Agency Conflicts and Cash: Estimates from a Dynamic Model," *Journal of Finance*, 69(5), 1883–1921.

Park, Y.-H., 2019, "Information in Yield Spread Trades," working paper, Board of Governors of the Federal Reserve System (US).

Penman, S. H., 1982, "Insider trading and the dissemination of firms' forecast information," *Journal of Business*, pp. 479–503.

Quadrini, V., 2004, "Investment and Liquidation in Renegotiation-Proof Contracts with Moral Hazard," *Journal of Monetary Economics*, 51, 713–751.

Rajan, R. G., and L. Zingales, 1998, "Financial Dependence and Growth," *American Economic Review*, 88(3), 559–586.

Rampini, A., and S. Viswanathan, 2010, "Collateral, Risk Management, and the Distribution of Debt Capacity," *Journal of Finance*, 65, 2293–2322.

Restuccia, D., and R. Rogerson, 2008, "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 11(4), 707–720.

Riddick, L. A., and T. M. Whited, 2009, "The Corporate Propensity to Save," *Journal of Finance*, 64(4), 1729–1766.

Sannikov, Y., 2008, "A Continuous-Time Version of the Principal-Agent Problem," *Review of Economic Studies*, 75(3), 957–984.

Shue, K., 2013, "Executive Networks and Firm Policies: Evidence from the Random Assignment of MBA Peers," *Review of Financial Studies*, 26(6), 1401–1442.

Sinha, P., and C. Gadarowski, 2010, "The efficacy of regulation fair disclosure," *Financial Review*, 45(2), 331–354.

Spear, S. E., and S. Srivastava, 1987, "On Repeated Model Hazard with Discounting," *Review of Economic Studies*, 54(4), 599–617.

Stohs, M. H., and D. C. Mauer, 1996, "The determinants of corporate debt maturity structure," *Journal of Business*, pp. 279–312.

Stoll, H. R., 1978, "The supply of dealer services in securities markets," *The Journal of Finance*, 33(4), 1133–1151.

Strebulaev, I. A., 2007, "Do tests of capital structure theory mean what they say?," *The Journal of Finance*, 62(4), 1747–1787.

Subrahmanyam, A., 1991, "Risk aversion, market liquidity, and price efficiency," *The Review of Financial Studies*, 4(3), 417–441.

Taylor, L. A., 2013, "CEO Wage Dynamics: Estimates from a Learning Model," *Journal of Financial Economics*, 108, 79–98.

Titman, S., and S. Tsyplakov, 2007, "A dynamic model of optimal capital structure," *Review of Finance*, 11(3), 401–451.

Titman, S., and R. Wessels, 1988, "The determinants of capital structure choice," *The Journal of Finance*, 43(1), 1–19.

Tong, J., and C. Ying, 2019, "A Dynamic-Agency Based Asset Pricing Theory with Production," Working Paper.

van Binsbergen, J. H., J. R. Graham, and J. Yang, 2010, "The cost of debt," *The Journal of Finance*, 65(6), 2089–2136.

Vayanos, D., 2001, "Strategic trading in a dynamic noisy market," *The Journal of Finance*, 56(1), 131–171.

Vissing-Jorgensen, A., 2020, "Central Banking with Many Voices: The Communications Arms Race," working paper, University of California Berkeley.

Wachter, J. A., and Y. Zhu, 2018, "The macroeconomic announcement premium," working paper, National Bureau of Economic Research.

Ward, C., 2019, "Agency in Intangibles," Working Paper.

Welch, I., 2004, "Capital structure and stock returns," *Journal of Political Economy*, 112(1), 106–131.

Whited, T. M., and G. Wu, 2006, "Financial Constraints Risk," *Review of Financial Studies*, 19(2), 531–559.

Ying, C., 2020, "Heterogeneous Beliefs and FOMC Announcements," *Available at SSRN 3286682*.

Zhu, J. Y., 2013, "Optimal Contracts with Shirking," *Review of Economic Studies*, 80(2), 812–839.

Zwiebel, J., 1996, "Dynamic Capital Structure under Managerial Entrenchment," *American Economic Review*, 86(5), 1197–.

Appendix A

Chapter 1 Appendices

The following appendices provide details of the proof in section 1.2, 1.3, and 1.4. Appendix A shows the related proof of the macroeconomic environment in section 1.2. Appendix B contains all the proofs for the benchmark economy that the insider is always informed. Appendix C provides the details for the economy that the market makers are uncertain about whether the insider is informed or not.

A Proof of the standard asset pricing framework

Proof of Lemma 1. The HJB equation for recursive utility satisfies

$$\bar{f}\left(C_t, V\left(\hat{m}_t, t, C_t\right)\right) + \mathcal{L}\left[\left(\hat{m}_t, t, C_t\right)\right] = 0.$$

Due to homogeneity, consider the value function of the form

$$V\left(\hat{m}_t, t, C_t\right) = \frac{1}{1-\gamma} H\left(\hat{m}_t, t\right) C_t^{1-\gamma},$$

where $H\left(\hat{m}_t, t\right)$ satisfies the following HJB equation:

$$\begin{aligned} 0 \quad = \quad & \frac{\rho}{1-\frac{1}{\psi}}\left(H\left(\hat{m}_t, t\right)^{-\frac{1-\frac{1}{\psi}}{1-\gamma}} - 1\right) + \left(\hat{m}_t - \frac{1}{2}\gamma\sigma_C^2\right) + \frac{1}{1-\gamma}\frac{H_t\left(\hat{m}_t, t\right)}{H\left(\hat{m}_t, t\right)} \\ & + \left[\frac{1}{1-\gamma}a_m\left(\bar{m} - \hat{m}_t\right) + q_t\right]\frac{H_m\left(\hat{m}_t, t\right)}{H\left(\hat{m}_t, t\right)} + \frac{1}{2}\frac{1}{1-\gamma}\frac{H_{mm}\left(\hat{m}_t, t\right)}{H\left(\hat{m}_t, t\right)}\left(\frac{q_t}{\sigma_C}\right)^2. \end{aligned} \tag{A1}$$

with the boundary condition that

$$H\left(\hat{m}_t^-, t^-\right) = \mathbb{E}\left[H\left(\hat{m}_t^+, t^+\right) | \hat{m}_t^-, q_t^-\right]. \tag{A2}$$

The state price process of recursive utility satisfies

$$\frac{d\pi_t}{\pi_t} = \frac{d\bar{f}_C\left(C_t, V_t\right)}{\bar{f}_C\left(C_t, V_t\right)} + \bar{f}_V\left(C_t, V_t\right)dt. \tag{A3}$$

Therefore, for $n = 1, 2, \cdots$, in the interior of $(nT, (n+1)T)$, the law of motion of the state price density, π_t satisfies the stochastic differential equation of the form:

$$d\pi_t = \pi_t\left[-r\left(\hat{m}, t\right)dt - \sigma_\pi\left(\hat{m}, t\right)d\tilde{B}_{C,t}\right],$$

where

$$r\left(\hat{m}, t\right) = \rho + \frac{1}{\psi}\hat{m} - \frac{1}{2}\gamma\left(1 + \frac{1}{\psi}\right)\sigma_C^2 - \frac{\gamma - \frac{1}{\psi}}{1-\gamma}\frac{H_m\left(\hat{m}_t, t\right)}{H\left(\hat{m}_t, t\right)}q\left(t\right) + \frac{1}{2}\frac{\frac{1}{\psi} - \gamma}{1-\gamma}\frac{1 - \frac{1}{\psi}}{1-\gamma}\left(\frac{H_m\left(\hat{m}_t, t\right)}{H\left(\hat{m}_t, t\right)}\right)^2\left(\frac{q\left(t\right)}{\sigma_C}\right)^2,$$

is the risk-free interest rate, and

$$\sigma_\pi\left(\hat{m}, t\right) = \gamma\sigma_C - \frac{\frac{1}{\psi} - \gamma}{1-\gamma}\frac{H_m\left(\hat{m}_t, t\right)}{H\left(\hat{m}_t, t\right)}\frac{q\left(t\right)}{\sigma_C},$$

is the market price of the Brownian motion risk.

Upon announcements, the stochastic discount factor for a small interval $\triangle$ is

$$\mathrm{SDF}_{t,t+\triangle} = e^{-\rho\triangle}\left(\frac{C_{t+\triangle}}{C_t}\right)^{-\frac{1}{\psi}}\left(\frac{W_{t+\triangle}}{E_t\left[W_{t+\triangle}^{1-\gamma}\right]^{\frac{1}{1-\gamma}}}\right)^{\frac{1}{\psi}-\gamma}, \tag{A4}$$

where

$$W_t = \left[\left(1-\gamma\right)V\left(\hat{m}, t, C\right)\right]^{\frac{1}{1-\gamma}}, \tag{A5}$$

which implies

$$\mathrm{SDF}_{t,t+\triangle} = e^{-\rho\triangle}\left(\frac{C_{t+\triangle}}{C_t}\right)^{-\frac{1}{\psi}}\left(\frac{\left(H\left(\hat{m}_{t+\Delta}, t+\triangle\right)C_{t+\Delta}^{1-\gamma}\right)^{\frac{1}{1-\gamma}}}{\mathbb{E}_t\left[H\left(\hat{m}_{t+\Delta}, t+\triangle\right)C_{t+\Delta}^{1-\gamma}\right]^{\frac{1}{1-\gamma}}}\right)^{\frac{1}{\psi}-\gamma}.$$

Therefore, when upon announcements $(t = nT)$ and $\triangle \to 0$,

$$\Lambda^*_{t,t+\triangle} = \frac{\left[H\left(\hat{m}_{t+\Delta}, t+\Delta\right)\right]^{\frac{\frac{1}{\psi}-\gamma}{1-\gamma}}}{\left[\mathbb{E}_t\left(H\left(\hat{m}_{t+\Delta}, t+\Delta\right)\right)\right]^{\frac{\frac{1}{\psi}-\gamma}{1-\gamma}}}. \tag{A6}$$

The term $\beta H\left(\hat{m}_t, t\right)^{-\frac{1-\frac{1}{\psi}}{1-\gamma}}$ is the consumption-wealth ratio. Consider the following log-linear expansion: $e^{\ln x} \approx e^{\ln\bar{x}} + e^{\ln\bar{x}}\left(\ln x - \ln\bar{x}\right)$,

$$\beta H\left(\hat{m}_t, t\right)^{-\frac{1-\frac{1}{\psi}}{1-\gamma}} \approx \kappa + \kappa\left[\ln\rho - \frac{1-\frac{1}{\psi}}{1-\gamma}\ln H\left(\hat{m}_t, t\right) - \ln\kappa\right],$$

where $\kappa=\beta H\left(\bar{m},t\right)^{-\frac{1-\frac{1}{\psi}}{1-\gamma}}$ is the consumption-wealth ratio when $\hat{m}_t$ is equal to its unconditional mean $\bar{m}$. Therefore, I can approximate $\beta H\left(\hat{m}_t,t\right)^{-\frac{1-\frac{1}{\psi}}{1-\gamma}}$ as

$$
\begin{aligned}
\frac{\rho}{1-\frac{1}{\psi}}\left[H\left(\hat{m}_t,t\right)^{1-\frac{1-\frac{1}{\psi}}{1-\gamma}}-1\right] &\approx \frac{1}{1-\frac{1}{\psi}}\left[\kappa+\kappa\left[\ln\rho-\frac{1-\frac{1}{\psi}}{1-\gamma}\ln H\left(\hat{m}_t,t\right)-\ln\kappa\right]-\rho\right] \\
&= -\frac{\kappa}{1-\gamma}\ln H\left(\hat{m}_t,t\right)+\xi_0,
\end{aligned}
$$

where I denote $\xi_0\triangleq\frac{1}{1-\frac{1}{\psi}}\left[\kappa-\rho-\kappa\left(\ln\kappa-\ln\rho\right)\right]$.

The HJB equation (A1) is written as

$$
\begin{aligned}
&\xi_0-\frac{\kappa}{1-\gamma}\ln H\left(\hat{m}_t,t\right)+\left(\hat{m}_t-\frac{1}{2}\gamma\sigma_C^2\right)+\frac{1}{1-\gamma}\frac{H_t\left(\hat{m}_t,t\right)}{H\left(\hat{m}_t,t\right)} \\
&+\left[\frac{1}{1-\gamma}a_m\left(\bar{m}-\hat{m}_t\right)+q_t\right]\frac{H_m\left(\hat{m}_t,t\right)}{H\left(\hat{m}_t,t\right)}+\frac{1}{2}\frac{1}{1-\gamma}\frac{H_{mm}\left(\hat{m}_t,t\right)}{H\left(\hat{m}_t,t\right)}\left(\frac{q_t}{\sigma_C}\right)^2=0.
\end{aligned}
\tag{A7}
$$

I guess $H\left(\hat{m}_t,t\right)$ is of the form

$$
H\left(\hat{m}_t,t\right)=e^{-\gamma^A\hat{m}_t+\mathcal{H}(t)}. \tag{A8}
$$

Using the method of undetermined coefficients, I have that

$$
\gamma^A = \frac{\gamma-1}{a_m+\kappa}, \tag{A9}
$$

$$
\mathcal{H}'\left(t\right) = \kappa\mathcal{H}\left(t\right)-f\left(t\right). \tag{A10}
$$

where $f\left(t\right)$ is defined as:

$$
f\left(t\right)=\frac{\left(1-\gamma\right)^2}{a_m+\kappa}q\left(t\right)+\frac{1}{2}\frac{\left(1-\gamma\right)^2}{\left(a_m+\kappa\right)^2}\frac{1}{\sigma_C^2}q^2\left(t\right)-\frac{1}{2}\gamma\left(1-\gamma\right)\sigma_C^2+a_m\bar{m}\frac{1-\gamma}{a_m+\kappa}+\xi_0.
$$

$\mathcal{H}\left(t\right)$ can be solved in closed form from equations (A10) and (A2). In order to solve for asset prices, I do not need the functional form $\mathcal{H}\left(t\right)$.

Note that this above approximation is *exact* if $\psi=1$, in which case

$$
\gamma^A=\frac{\gamma-1}{a_m+\rho}. \tag{A11}
$$

Besides, from equations (A6) and (A9), it is straightforward to show the A-SDF is countercyclical if and only if the agent has early resolution of uncertainty, i.e., $\gamma>\frac{1}{\psi}$, which is equivalent to $\gamma^A>0$ when $\psi=1$. ■

Lemma 5. *Under the assumption that the aggregate consumption does not change in the 24-hour window before announcements, at $t=nT-1$, the agent has a prior that the expected growth rate upon announcements $\hat{m}_{nT}$ is normally distributed $N\left(\hat{m}_{nT-1},\Delta Q\right)$ where $\Delta Q=q_{nT-1}-q_{nT}$.*

Proof. At announcements $t = nT$, the agent updates her belief using Bayes' rule:

$$\hat{m}_{nT} = q_{nT}\left[\frac{1}{\sigma_s^2}s_n + \frac{1}{q_{nT-1}}\hat{m}_{nT-1}\right], \quad \frac{1}{q_{nT}} = \frac{1}{\sigma_s^2} + \frac{1}{q_{nT-1}}, \tag{A12}$$

which implies

$$\mathbb{E}_{nT-1}\left[\hat{m}_{nT}\right] = \hat{m}_{nT-1}, \quad Var_{nT-1}\left[\hat{m}_{nT}\right] = \left(\frac{q_{nT}}{\sigma_s^2}\right)^2\left(q_{nT-1} + \sigma_s^2\right) = q_{nT-1} - q_{nT}. \tag{A13}$$

Therefore, at $t = nT-1$, m_{nT} is normally distributed $N\left(\hat{m}_{nT-1}, \Delta Q\right)$ where $\Delta Q = q_{nT-1} - q_{nT}$.
■

B Proof of Theorem 1

The proof is in several steps.

B.1 Step 1: Market Maker's Updating

First, I establish that if the market makers conjecture that the insider's trading strategy follows equation (1.15), then the price dynamics equation (1.20) satisfies the market makers' break-even pricing rule given in equation (1.10).

Proof of Lemma 2. The conjectured trading strategy (1.15) implies that

$$\begin{aligned}
\theta_t &= \frac{\log\left[A\left(\hat{m}_{nT}, nT\right)\right] - \mu_P + \gamma^A\beta\Delta Q}{\left(nT-t\right)\lambda} - \frac{Y_t}{nT-t} \\
&= \frac{\frac{\beta-\gamma^A}{\beta}\left\{\log\left[A\left(\hat{m}_{nT}, nT\right)\right] - \mu_P + \gamma^A\beta\Delta Q\right\}}{\left(nT-t\right)\left(\frac{\beta-\gamma^A}{\beta}\lambda\right)} - \frac{Y_t}{nT-t} \\
&= \frac{\left(\log\left[H\left(\hat{m}_{nT}, nT\right)A\left(\hat{m}_{nT}, nT\right)\right] - \left(\mu_V - \frac{\gamma^A\left(\beta-\gamma^A\right)}{\beta^2}\sigma_v^2\right)\right)/\left(\frac{\beta-\gamma^A}{\beta}\lambda\right) - Y_t}{nT-t},
\end{aligned} \tag{A14}$$

where the last equality comes from $H\left(\hat{m}_{nT}, nT\right) = e^{-\gamma^A\hat{m}_{nT} + \mathcal{H}(nT)}$ and $A\left(\hat{m}_{nT}, nT\right) = e^{\beta\hat{m}_{nT} + N(nT)}$. Here $\mu_P = \beta\hat{m}_{nT-1} + N\left(nT\right)$ and $\mu_V = \left(\beta - \gamma^A\right)\hat{m}_{nT-1} + \mathcal{H}\left(nT\right) + N\left(nT\right)$.

Therefore, the aggregate trading volume follows

$$dY_t = \theta_t dt + dZ_t = \frac{\left(\log\left[H\left(\hat{m}_{nT}, nT\right)A\left(\hat{m}_{nT}, nT\right)\right] - \left(\mu_V - \frac{\gamma^A\left(\beta-\gamma^A\right)}{\beta^2}\sigma_v^2\right)\right)/\left(\frac{\beta-\gamma^A}{\beta}\lambda\right) - Y_t}{nT-t}dt + \sigma_z dB_t, \tag{A15}$$

where $dZ_t = \sigma_z dB_t$ and $Y_{nT-1} = 0$.

Now let me define the observation and innovation process. Set $Y^*_{nT-1} = 0$ and

$$dY^*_t = \frac{1}{\sigma_z}\left(dY_t + \frac{\left(\mu_V - \frac{\gamma^A(\beta-\gamma^A)}{\beta^2}\sigma_v^2\right)/\left(\frac{\beta-\gamma^A}{\beta}\lambda\right) + Y_t}{nT-t}dt\right)$$
$$= \frac{\log\left[H\left(\hat{m}_{nT}, nT\right)A\left(\hat{m}_{nT}, nT\right)\right]}{\frac{\beta-\gamma^A}{\beta}\sigma_v\left(nT-t\right)}dt + dB_t,$$

where the last equality comes from $\lambda = \frac{\sigma_v}{\sigma_z}$. Because Y_t are observable to market makers, Y^* is also observable. The corresponding innovation process is given by

$$dB^*_t = \frac{\log\left[H\left(\hat{m}_{nT}, nT\right)A\left(\hat{m}_{nT}, nT\right)\right] - \hat{v}_t}{\frac{\beta-\gamma^A}{\beta}\sigma_v\left(nT-t\right)}dt + dB_t$$

where

$$\hat{v}_t = \mathbb{E}\left[\log\left[H\left(\hat{m}_{nT}, nT\right)A\left(\hat{m}_{nT}, nT\right)\right]|\mathcal{F}^Y_t\right]. \tag{A16}$$

The Kalman-filter equation implies

$$d\hat{v}_t = \frac{\sum_{v,t}}{\frac{\beta-\gamma^A}{\beta}\sigma_v\left(nT-t\right)}dB^*_t, \tag{A17}$$

where

$$\sum\nolimits_{v,t} = Var\left[\log H\left(\hat{m}_{nT}, nT\right)v\left(\hat{m}_{nT}, nT\right)|\mathcal{F}^Y_t\right], \tag{A18}$$

is the conditional variance of $\log\left[H\left(\hat{m}_{nT}, nT\right)A\left(\hat{m}_{nT}, nT\right)\right]$ given market marker's information (on the filtration $\mathcal{F}^Y_t$). The Kalman-filter equation also implies the dynamics of the posterior variance:

$$\frac{1}{\sum_{v,t}} = \frac{1}{\sum_{v,0}} + \int_{nT-1}^{t}\frac{1}{\left(nT-s\right)^2\left(\frac{\beta-\gamma^A}{\beta}\right)^2\sigma_v^2}ds = \frac{1}{\left(\frac{\beta-\gamma^A}{\beta}\right)^2\sigma_v^2} + \frac{t-\left(nT-1\right)}{\left(nT-t\right)\left(\frac{\beta-\gamma^A}{\beta}\right)^2\sigma_v^2} = \frac{1}{\left(nT-t\right)\left(\frac{\beta-\gamma^A}{\beta}\right)^2\sigma_v^2}, \tag{A19}$$

which implies

$$\sum\nolimits_{v,t} = \left(\frac{\beta-\gamma^A}{\beta}\right)^2\sigma_v^2\left(nT-t\right). \tag{A20}$$

Thus, the filtering equation (A17) is

$$d\hat{v}_t = \frac{\beta-\gamma^A}{\beta}\sigma_v dB^*_t = \frac{\log\left[H\left(\hat{m}_{nT}, nT\right)A\left(\hat{m}_{nT}, nT\right)\right] - \hat{v}_t}{nT-t}dt + \frac{\beta-\gamma^A}{\beta}\sigma_v dB_t. \tag{A21}$$

Define an adjusted order flow $\hat{Y}_t$ as

$$\hat{Y}_t \equiv Y_t - \int_{nT-1}^{t}\left(\frac{\gamma^A\beta\Delta Q}{\lambda}\right)ds = Y_t - \frac{\gamma^A\beta\Delta Q}{\lambda}\left[t-\left(nT-1\right)\right]. \tag{A22}$$

From the aggregate trading volume (A15), the adjusted order flow follows

$$
\begin{aligned}
d\hat{Y}_t &= dY_t - \frac{\gamma^A \beta \Delta Q}{\lambda} dt \\
&= \frac{\left(\log\left[H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right)\right] - \mu_V\right) / \left(\frac{\beta-\gamma^A}{\beta}\lambda\right) - \hat{Y}_t}{nT - t} dt + \sigma_z dB_t. \qquad \text{(A23)}
\end{aligned}
$$

This implies

$$
\frac{\beta - \gamma^A}{\beta}\lambda d\hat{Y}_t = \frac{\log\left[H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right)\right] - \left[\mu_V + \frac{\beta-\gamma^A}{\beta}\lambda \hat{Y}_t\right]}{nT - t} dt + \frac{\beta - \gamma^A}{\beta}\sigma_v dB_t. \quad \text{(A24)}
$$

Since $\hat{v}_{nT-1} = \mathbb{E}\left[\log H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right) | \mathcal{F}^Y_{nT-1}\right] = \mu_V$ and $\hat{Y}_{nT-1} = 0$, combining (A21) and (A24) gives,

$$
d\hat{v}_t = \frac{\beta - \gamma^A}{\beta}\sigma_v dB^*_t = \frac{\beta - \gamma^A}{\beta}\lambda d\hat{Y}_t = \frac{\beta - \gamma^A}{\beta}\lambda\left[dY_t - \frac{\gamma^A \beta \Delta Q}{\lambda} dt\right], \qquad \text{(A25)}
$$

where the last equality holds due to the definition of the adjusted order flow in equation (A22). From the filtering theory, B^*_t is a standard Brownian Motion with respect to market makers' filtration. Therefore, the adjusted order flow $\hat{Y}_t$ is a Brownian Motion with instant variance σ_z^2 under $\mathcal{F}^Y_t$. This also implies

$$
\mathbb{E}\left[\theta_t | \mathcal{F}^Y_t\right] = \frac{\gamma^A \beta \Delta Q}{\lambda}, \qquad \text{(A26)}
$$

is the market makers' expectation of the insider's order rate, which is strictly positive when the market makers are risk compensated.

The market makers' prior about $\log\left[H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right)\right]$ at time $nT - 1$ is represented by a normal distribution. The Kalman filter implies the posterior distribution of $\log\left[H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right)\right]$ under $\mathcal{F}^Y_t$ is also Gaussian, which is summarized by the posterior mean $\hat{v}_t$ and the posterior variance $\sum_{v,t}$. Therefore, $\forall t \in [nT - 1, nT]$, the market makers' estimation of $H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right)$ is

$$
\begin{aligned}
V_t &= \mathbb{E}\left[H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right) | \mathcal{F}^Y_t\right] \\
&= \mathbb{E}\left[e^{\log H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right)} | \mathcal{F}^Y_t\right] \\
&= e^{\hat{v}_t + \frac{1}{2}\sum_{v,t}} = e^{\hat{v}_t + \frac{1}{2}(nT-t)\left(\frac{\beta-\gamma^A}{\beta}\right)^2 \sigma_v^2}. \qquad \text{(A27)}
\end{aligned}
$$

Applying Ito's Lemma, from equation (A25), I find

$$
\begin{aligned}
\frac{dV_t}{V_t} &= \frac{1}{V_t}\left[V_t d\hat{v}_t + \frac{1}{2}V_t\left(d\hat{v}_t\right)^2 - \frac{1}{2}\sigma_v^2 V_t dt\right] \\
&= d\hat{v}_t = \frac{\beta - \gamma^A}{\beta}\lambda d\hat{Y}_t. \qquad \text{(A28)}
\end{aligned}
$$

Similarly, I define Λ_t^* as the posterior mean of $\log H(\hat{m}_{nT}, nT)$ under market makers' information:

$$\begin{aligned}
\Lambda_t^* &= \mathbb{E}\left[\log H(\hat{m}_{nT}, nT) | \mathcal{F}_t^Y\right] \\
&= -\gamma^A \mathbb{E}\left[\hat{m}_{nT} | \mathcal{F}_t^Y\right] + \mathcal{H}(nT) \\
&= \frac{-\gamma^A}{\beta - \gamma^A} \mathbb{E}\left[\left(\beta - \gamma^A\right) \hat{m}_{nT} | \mathcal{F}_t^Y\right] + \mathcal{H}(nT) \\
&= \frac{-\gamma^A}{\beta - \gamma^A} \mathbb{E}\left[\log H(\hat{m}_{nT}, nT) A(\hat{m}_{nT}, nT) | \mathcal{F}_t^Y\right] + \frac{\beta \mathcal{H}(nT) + \gamma^A N(nT)}{\beta - \gamma^A} \\
&= \frac{-\gamma^A}{\beta - \gamma^A} \hat{v}_t + \frac{\beta \mathcal{H}(nT) + \gamma^A N(nT)}{\beta - \gamma^A} \qquad \text{(A29)}
\end{aligned}$$

where the last equality holds due to equation (A16). It implies

$$d\Lambda_t^* = \frac{-\gamma^A}{\beta - \gamma^A} d\hat{v}_t,$$

with $\Lambda_{nT-1}^* = -\gamma^A \hat{m}_{nT-1} + \mathcal{H}(nT)$. Therefore, the posterior variance of $\log H(\hat{m}_{nT}, nT)$ under the market makers' information is,

$$\Sigma_{\Lambda^*,t} = Var\left[\log\left[H(\hat{m}_{nT}, nT)\right] | \mathcal{F}_t^Y\right] = \left(\frac{\gamma^A}{\beta - \gamma^A}\right)^2 \Sigma_{v,t} = \left(\frac{\gamma^A}{\beta}\right)^2 (nT - t)\,\sigma_v^2.$$

The Kalman filter implies the posterior distribution of $\log\left[H(\hat{m}_{nT}, nT)\right]$ under $\mathcal{F}_t^Y$ is also Gaussian, which is summarized by the posterior mean Λ_t^* and the posterior variance $\Sigma_{\Lambda^*,t}$. Therefore, $\forall t \in [nT-1,\, nT]$, the market makers' estimation of $H(\hat{m}_{nT}, nT)$ is

$$\Lambda_t = \mathbb{E}\left[H(\hat{m}_{nT}, nT) | \mathcal{F}_t^Y\right] = \mathbb{E}\left[e^{\log H(\hat{m}_{nT}, nT)} | \mathcal{F}_t^Y\right] = e^{\Lambda_t^* + \frac{1}{2}\Sigma_{\Lambda^*,t}} = e^{\Lambda_t^* + \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 (nT-t)\sigma_v^2}.$$

From Ito's Lemma,

$$\begin{aligned}
\frac{d\Lambda_t}{\Lambda_t} &= \frac{1}{\Lambda_t}\left[\Lambda_t d\Lambda_t^* + \frac{1}{2}\Lambda_t \left(d\Lambda_t^*\right)^2 - \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2 \Lambda_t dt\right] \\
&= d\Lambda_t^* = \frac{-\gamma^A}{\beta - \gamma^A} d\hat{v}_t = -\frac{\gamma^A}{\beta} \lambda d\hat{Y}_t. \qquad \text{(A30)}
\end{aligned}$$

Therefore, both V_t and Λ_t are functions of the adjusted order flow $\hat{Y}_t$. From the definition of price dynamics in equation (1.14),

$$P_t = \frac{\mathbb{E}\left[H(\hat{m}_{nT}, nT) A(\hat{m}_{nT}, nT) | \mathcal{F}_t^Y\right]}{\mathbb{E}\left[H(\hat{m}_{nT}, nT) | \mathcal{F}_t^Y\right]} = \frac{V\left(t, \hat{Y}_t\right)}{\Lambda\left(t, \hat{Y}_t\right)},$$

the equilibrium pricing rule is also a function of the adjusted order flow, i.e., $P\left(t, \hat{Y}_t\right)$.

I apply Ito's Lemma to V_t,

$$\frac{dV_t}{V_t} = \frac{d\left(P_t \Lambda_t\right)}{P_t \Lambda_t} = \frac{dP_t}{P_t} + \frac{d\Lambda_t}{\Lambda_t} + \frac{dP_t}{P_t}\frac{d\Lambda_t}{\Lambda_t}. \tag{A31}$$

From equations (A28) and (A30), I find

$$\frac{dP\left(t, \hat{Y}_t\right)}{P\left(t, \hat{Y}_t\right)} = \lambda d\hat{Y}_t + \gamma^A \beta \Delta Q dt, \text{ with } P_{nT-1} = e^{\mu_P - \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2 + \frac{1}{2}\left(\frac{\beta - \gamma^A}{\beta}\right)^2 \sigma_v^2}. \tag{A32}$$

Furthermore, from equation (A23), the process $\hat{Y}_t$ is a Brownian bridge with instantaneous variance σ_z^2 with respect to the insider's filtration, terminating at $\left(\log\left[H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right)\right] - \mu_V\right) / \left(\frac{\beta - \gamma^A}{\beta}\lambda\right)$ (Karatzas and Shreve (1987)). It satisfies $\hat{Y}_t \to \left(\log\left[H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right)\right] - \mu_V\right) / \left(\frac{\beta - \gamma^A}{\beta}\lambda\right)$ with probability 1 as $t \to nT$.[1] This implies the equilibrium price in equation (A32) satisfies :

$$\begin{aligned}
\log P_t &= \log P_{nT-1} + \lambda \hat{Y}_t - \left[\frac{1}{2}\sigma_v^2 - \gamma^A \beta \Delta Q\right]\left(t - (nT-1)\right) \qquad \text{(A33)} \\
&= \log P_{nT-1} + \lambda \hat{Y}_t - \frac{1}{2}\frac{\beta - 2\gamma^A}{\beta}\sigma_v^2\left(t - (nT-1)\right) \\
&\to \beta \hat{m}_{nT-1} - \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2 + N(nT) + \frac{1}{2}\left(\frac{\beta - \gamma^A}{\beta}\right)^2 \sigma_v^2 + \beta\left(\hat{m}_{nT} - \hat{m}_{nT-1}\right) - \frac{1}{2}\frac{\beta - 2\gamma^A}{\beta}\sigma_v^2 \\
&\to \beta \hat{m}_{nT} + N(nT) = \log A\left(\hat{m}_{nT}, nT\right).
\end{aligned}$$

almost surely as $t \to nT$ from the insider's information. This is equivalent to $P_t \to A\left(\hat{m}_{nT}, nT\right)$ with probability 1 as $t \to nT$ under the insider's filtration.

■

B.2 Step 2: Insider's Optimal Strategy

Second, I capture the insider's optimal trading strategy when the equilibrium pricing rule is a function of the adjusted order flow, i.e., $P\left(t, \hat{Y}_t\right)$.

Proof of Lemma 3. By Thereom 7.6 in Chapter 5 of Karatzas and Shreve (1991) (Feynman-Kac representation), the value function J defined in equation (1.26), is a unique solution to the Bellman equation (1.24) with the terminal condition $J\left(nT, y, A\left(\hat{m}_{nT}, nT\right)\right) = j\left(y, A\left(\hat{m}_{nT}, nT\right)\right)$.

[1] The distribution of a Brownian bridge is the same as a Brownian motion conditional on the terminal value being known. $\left(\log\left[H\left(\hat{m}_{nT}, nT\right) A\left(\hat{m}_{nT}, nT\right)\right] - \mu_V\right) / \left(\frac{\beta - \gamma^A}{\beta}\lambda\right)$ is the terminal value of $\hat{Y}_t$, which is normally distributed with mean zero and variance σ_z^2 and is independent of Z. Hence, the distribution of $\hat{Y}_t$, unconditional on the terminal value or Z (i.e., from the market makers' filtration), are the distribution of a Brownian motion with variance σ_z^2. This is consistent with what I get from the filtering theory in equation (A25).

Taking the derivative under the expectation operator yields[2]

$$\begin{aligned}J_y\left(t,y,A\left(\hat{m}_{nT},nT\right)\right)&=\mathbb{E}\left[j_y\left(y+\omega_{nT}-\omega_t,A\left(\hat{m}_{nT},nT\right)\right)\right]\\&=\mathbb{E}\left[g\left(y+\omega_{nT}-\omega_t\right)\right]-A\left(\hat{m}_{nT},nT\right)\\&=P\left(t,y\right)-A\left(\hat{m}_{nT},nT\right),\end{aligned}$$

which shows $J\left(t,y,A\left(\hat{m}_{nT},nT\right)\right)$ also satisfy equation (1.23) with $P\left(t,y\right)$ as defined by (1.25).

■

Proof of Lemma 4. For any trading strategy θ_t, apply Ito's Lemma to the value function,

$$\begin{aligned}J\left(nT,\hat{Y}_{nT},A\left(\hat{m}_{nT},nT\right)\right)&=J\left(nT-1,\hat{Y}_{nT-1},A\left(\hat{m}_{nT},nT\right)\right)+\int_{nT-1}^{nT}\left\{J_tdt+J_yd\hat{Y}_t+\frac{1}{2}J_{yy}\left(d\hat{Y}_t\right)^2\right\}\\&=J\left(nT-1,\hat{Y}_{nT-1},A\left(\hat{m}_{nT},nT\right)\right)+\int_{nT-1}^{nT}\left\{J_tdt+J_y\left(\left[\theta_t-\hat{\theta}_t\right]dt+dZ_t\right)+\frac{1}{2}\sigma_z^2J_{yy}\right\}\\&=J\left(nT-1,\hat{Y}_{nT-1},A\left(\hat{m}_{nT},nT\right)\right)-\int_{nT-1}^{nT}\left(A\left(\hat{m}_{nT},nT\right)-P\left(t,\hat{Y}_t\right)\right)\left(\theta_tdt+dZ_t\right),\end{aligned}$$

where I use equations (1.23) and (1.24). I can rearrange this as

$$\begin{aligned}\int_{nT-1}^{nT}\left(A\left(\hat{m}_{nT},nT\right)-P\left(t,\hat{Y}_t\right)\right)\theta_tdt&=J\left(nT-1,\hat{Y}_{nT-1},A\left(\hat{m}_{nT},nT\right)\right)-J\left(nT,\hat{Y}_{nT},A\left(\hat{m}_{nT},nT\right)\right)\\&\quad-\int_{nT-1}^{nT}\left(A\left(\hat{m}_{nT},nT\right)-P\left(t,\hat{Y}_t\right)\right)dZ_t\end{aligned}$$

The left-hand side is the profit of the insider, and the right-hand side is bounded above by

$$J\left(nT-1,\hat{Y}_{nT-1},A\left(\hat{m}_{nT},nT\right)\right)-\int_{nT-1}^{nT}\left(A\left(\hat{m}_{nT},nT\right)-P\left(t,\hat{Y}_t\right)\right)dZ_t \tag{A34}$$

due to the nonnegativity of $J\left(nT,\hat{Y}_{nT},A\left(\hat{m}_{nT},nT\right)\right)$ in equation (1.26). The no-double-strategies condition

$$E\int_{nT-1}^{nT}P_t^2dt<\infty$$

implies that the stochastic integral in (A34) has a zero expectation. Therefore,

$$E\int_{nT-1}^{nT}\left\{\left[A\left(\hat{m}_{nT},nT\right)P\left(t,\hat{Y}_t\right)\right]\theta_tdt\right\}t\leq J\left(nT-1,P_{nT-1},A\left(\hat{m}_{nT},nT\right)\right),$$

with equality if and only if $\hat{Y}_{nT}=g^{-1}\left(A\left(\hat{m}_{nT},nT\right)\right)$, which is equivalent to $P\left(nT,\hat{Y}_{nT}\right)=A\left(\hat{m}_{nT},nT\right)$ from equation (1.25). Thus, $J\left(nT-1,\hat{Y}_{nT-1},A\left(\hat{m}_{nT},nT\right)\right)$ is an upper bound

[2] The proof that the derivative of the right-hand side of (1.26) can be taken under the expectation operator is similar to Back (1992).

on the insider's expected profit, conditional on the termination value $A(\hat{m}_{nT}, nT)$, and the upper bound is realized - and the corresponding strategy is consequently optimal - if and only if $P\left(nT, \hat{Y}_{nT}\right) = A(\hat{m}_{nT}, nT)$.

■

Having established these results, finally, I show that the conjectured rule by the market makers is indeed consistent with the insider's optimal choice, as stated in Theorem 1.

Proof of Theorem 1. Since $\hat{Y}_{nT} = g^{-1}(A(\hat{m}_{nT}, nT))$ a.s., for any scalar a, the probability, given the market makers' information at time $nT-1$, that $\hat{Y}_{nT} \leq a$ is $F(g(A(\hat{m}_{nT}, nT)))$ where F is the distribution function of $A(\hat{m}_{nT}, nT)$. According to Lemma , the distribution function of $\hat{Y}_{nT}$, given the market makers' information at time 0, is normal distribution with mean zero and variance σ_z^2 and I denote it as N. Therefore, $N = F \circ g$, implying $g = F^{-1} \circ N$. When $\log A(\hat{m}_{nT}, nT)$ is normally distributed with mean $\beta \hat{m}_{nT-1} + N(nT)$ and variance $\beta^2 [q_{nT-1} - q_{nT}] = \sigma_v^2$. Set $g(y) = F^{-1}(N(y))$:

$$F(g(y)) = N^* \left(\frac{\log g(y) - [\beta \hat{m}_{nT-1} + N(nT)]}{\sigma_v} \right) = N^* \left(\frac{\hat{Y}_t}{\sigma_z} \right),$$

so

$$g(y) = \exp(\beta \hat{m}_{nT-1} + N(nT) + \lambda y), \tag{A35}$$

where $\lambda = \frac{\sigma_v}{\sigma_z}$ and $g(y)$ is a increasing function in y since $\lambda > 0$. From the conjectured trading strategy in equation (1.15), $\hat{\theta}_t \equiv \mathbb{E}\left[\theta_t | \mathcal{F}_t^Y\right] = \frac{\gamma^A \beta \Delta Q}{\lambda}$. It implies

$$\begin{aligned}
P\left(t, \hat{Y}_t\right) &= \mathbb{E}\left[g\left(\hat{Y}_t + \omega_{nT} - \omega_t\right)\right] \\
&= \mathbb{E}\left[\exp\left(\beta \hat{m}_{nT-1} + N(nT) + \lambda\left(\hat{Y}_t + Z_{nT} - Z_t - \frac{\gamma^A \beta \Delta Q}{\lambda}(nT - t)\right)\right)\right] \\
&= \exp\left(\beta \hat{m}_{nT-1} + N(nT) + \lambda \hat{Y}_t + \frac{1}{2}\sigma_v^2 (nT - t) - \gamma^A \beta \Delta Q (nT - t)\right) \\
&= \exp\left(\log P_{nT-1} + \lambda \hat{Y}_t - \left[\frac{1}{2}\sigma_v^2 - \gamma^A \beta \Delta Q\right](t - (nT - 1))\right)
\end{aligned} \tag{A36}$$

where $P_{nT-1} = e^{\beta \hat{m}_{nT-1} - \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2 + N(nT) + \frac{1}{2}\left(\frac{\beta - \gamma^A}{\beta}\right)^2 \sigma_v^2}$, which is exactly equation (A33). Therefore, the pricing function in equation (A35) implies the price dynamics follow equation (A32).

Equation (1.25) implies that $P(t, \omega_t)$ is a martingale under the filtration generated by ω. This implies the price dynamics and the expected trading volume $\hat{\theta}(t)$ with respect to $\mathcal{F}_t$ must satisfy

$$P_t - \hat{\theta}(t) P_y + \frac{1}{2}\sigma_z^2 P_{yy} = 0. \tag{A37}$$

It's very straightforward to show this pricing rule in equation (A36) satisfies the above property.

Besides, In the proof of Lemma 2, I have already shown that the trading strategy in (1.28) implies $P\left(t,\hat{Y}_t\right) \to A\left(\hat{m}_{nT}, nT\right)$ with probability 1 as $t \to nT$. It follows that the strategy (1.28) is optimal. Therefore, $\{P_t, \theta_t\}$ in equations (1.27) and (1.28) is an equilibrium.

Combining equations (1.26) and (A35), the maximized expected profit of the insider is

$$J\left(t, P\left(t,\hat{Y}_t\right), A\left(\hat{m}_{nT}, nT\right)\right) = \frac{1}{2}\sigma_v\sigma_z\left(nT-t\right)A\left(\hat{m}_{nT}, nT\right)$$
$$+\frac{P\left(t,\hat{Y}_t\right) - A\left(\hat{m}_{nT}, nT\right) + A\left(\hat{m}_{nT}, nT\right)\left[\log A\left(\hat{m}_{nT}, nT\right) - \log P\left(t,\hat{Y}_t\right)\right]}{\lambda}. \tag{A38}$$

As in Back (1992), I explicityly indicate the conditional expectation at time t given the market makers' information by $E^M[\cdot]$ and the conditional expectation given the insider's information by $E^I[\cdot]$. Given equation (A35), the pricing rule in equation (1.25) yields

$$\begin{aligned}
P\left(t, Z_t\right) &= E^I\left[g\left(Z_t + \omega_{nT} - \omega_t\right)|Z_t\right] \\
&\overset{(A35)}{=} E^I\left[\exp\left(\beta\hat{m}_{nT-1} + N\left(nT\right) + \lambda\left(Z_{nT} - \frac{\gamma^A\beta\Delta Q}{\lambda}\left(nT-t\right)\right)\right)|Z_t\right] \\
&= E^I\left[P\left(nT, Z_{nT}\right)|Z_t\right]\exp\left(-\gamma^A\beta\Delta Q\left(nT-t\right)\right),
\end{aligned} \tag{A39}$$

where the last equality comes from equation (A36) when $t = nT$:

$$\begin{aligned}
P\left(nT, Z_{nT}\right) &= \exp\left(\log P_{nT-1} + \lambda Z_{nT} - \left[\frac{1}{2}\sigma_v^2 - \gamma^A\beta\Delta Q\right]\right) \\
&= \exp\left(\beta\hat{m}_{nT-1} + N\left(nT\right) + \lambda Z_{nT}\right).
\end{aligned}$$

Rearrange equation (A39), I find

$$P\left(t, Z_t\right)\exp\left(-\gamma^A\beta\Delta Q\left(t-\left(nT-1\right)\right)\right) = E^I\left[P\left(nT, Z_{nT}\right)|Z_t\right]\exp\left(-\gamma^A\beta\Delta Q\right),$$

which implies $P\left(t, Z_t\right)\exp\left(-\gamma^A\beta\Delta Q\left(t-\left(nT-1\right)\right)\right)$ is a martingale under the insider's information set. Since the distribution of Z_t with respect to the insider's information is the same as the distribution of $\hat{Y}_t$ with respect to the market makers' information,

$$\begin{aligned}
P\left(t,\hat{Y}_t\right)\exp\left(-\gamma^A\beta\Delta Q\left(t-\left(nT-1\right)\right)\right) &= E^M\left[P\left(nT,\hat{Y}_{nT}\right)|\hat{Y}_t\right]\exp\left(-\gamma^A\beta\Delta Q\right) \\
&= E^M\left[P\left(nT,\hat{Y}_{nT}\right)|\left(\hat{Y}_s\right)_{s\le t}\right]\exp\left(-\gamma^A\beta\Delta Q\right)
\end{aligned}$$

where the last equality using the Markov property of a Brownian motion. This implies

$$P\left(t,\hat{Y}_t\right)\exp\left(-\gamma^A\beta\Delta Q\left(t-\left(nT-1\right)\right)\right)$$

is a martingale under the insider's information set. This is equivalent to say $P\left(t,\hat{Y}_t\right)$ is a submartingale with a deterministic growth rate $\gamma^A\beta\Delta Q$ per unit of time since both γ^A and β

are strictly positive. Similar argument applies to the price-response function $P_{\hat{Y}}\left(t,\hat{Y}\right)$, which is also a submartingale with a deterministic growth rate $\gamma^{A}\beta\Delta Q$ per unit of time.

Therefore, the unconditional expected return for any $t\in[nT-1,\,nT]$ is

$$\log\mathbb{E}\left[\frac{P_t}{P_{nT-1}}|\mathcal{F}^{Y}_{nT-1}\right]=\gamma^{A}\beta\Delta Q\left(t-(nT-1)\right)=\gamma^{A}\beta\Delta Q\left(t-(nT-1)\right), \tag{A40}$$

which implies the expected pre-FOMC announcement drift grows at a constant rate $\gamma^{A}\beta\Delta Q$. ■

Next, I prove the properties of the equilibrium in Theorem 1.

Proof of Proposition 1. In the meantime, the posterior variance of $\log P_{nT}$ at time $t\in[nT-1,\,nT]$ is

$$\begin{aligned}
Var\left[\log P_{nT}|\mathcal{F}^{Y}_{t}\right] &= Var\left[\log A\left(\hat{m}_{nT},nT\right)|\mathcal{F}^{Y}_{t}\right]\\
&= Var\left\{\log\left[H\left(\hat{m}_{nT},nT\right)A\left(\hat{m}_{nT},nT\right)\right]-\log\left[H\left(\hat{m}_{nT},nT\right)\right]|\mathcal{F}^{Y}_{t}\right\}\\
&= Var\left\{\log\left[H\left(\hat{m}_{nT},nT\right)A\left(\hat{m}_{nT},nT\right)\right]|\mathcal{F}^{Y}_{t}\right\}+Var\left\{\log\left[H\left(\hat{m}_{nT},nT\right)\right]|\mathcal{F}^{Y}_{t}\right\}\\
&\quad -2Cov\left(\log\left[H\left(\hat{m}_{nT},nT\right)A\left(\hat{m}_{nT},nT\right)\right],\log\left[H\left(\hat{m}_{nT},nT\right)\right]|\mathcal{F}^{Y}_{t}\right)\\
&= \left(\frac{\beta-\gamma^{A}}{\beta}\right)^{2}\sigma_v^2\left(nT-t\right)-2\frac{-\gamma^{A}\left(\beta-\gamma^{A}\right)}{\beta^2}\sigma_v^2\left(nT-t\right)+\left(\frac{\gamma^{A}}{\beta}\right)^{2}\sigma_v^2\left(nT-t\right)\\
&= \sigma_v^2\left(nT-t\right)=\beta^2\Delta Q\left(nT-t\right).
\end{aligned}$$

Therefore, the reduction of the uncertainty at time t comparing to $nT-1$ is

$$Var\left[\log P_{nT}|\mathcal{F}^{Y}_{t}\right]-Var\left[\log P_{nT}|\mathcal{F}^{Y}_{nT-1}\right]=\beta^2\Delta Q\left[(nT-1)-t\right],$$

which implies the uncertainty reduces at a constant rate $\beta^2\Delta Q$ per unit of time. ■

To prove Proposition 2, I first characterize the equilibrium when the market makers are risk-neutral, i.e., the original Kyle model under $\gamma^{A}_{\text{Kyle}}=0$.

Lemma 6. *When the market makers are risk-neutral, $\forall t\in[nT-1,\,nT]$, there exists an equilibrium where the price process P^{Kyle}_t and optimal strategy of the insider θ^{Kyle}_t have dynamics,*

$$\frac{dP^{Kyle}\left(t,Y^{Kyle}_t\right)}{P^{Kyle}\left(t,Y^{Kyle}_t\right)} = \lambda dY^{Kyle}_t \quad \text{with } P^{Kyle}_{nT-1}=e^{\mu_P+\frac{1}{2}\sigma_v^2}, \tag{A41}$$

$$\theta^{Kyle}\left(t,Y^{Kyle}_t\right) = \frac{\log\left[A\left(\hat{m}_{nT},nT\right)\right]-\mu_P}{\left(nT-t\right)\lambda}-\frac{Y^{Kyle}_t}{nT-t}, \tag{A42}$$

where μ_P, σ_v, and λ are defined in Lemma 2. Here the dynamics of aggregate order flow Y^{Kyle} is the sum of the insider's demand the liquidity traders' demand:

$$dY^{Kyle}_t=\theta^{Kyle}_t dt+dZ_t.$$

Proof. See Back (1992) or let $\gamma^A = 0$ in Theorem 1. ■

Proof of Proposition 2. It is a straightforward calculation from the equations (1.27), (1.28), (A41) and (A42). ■

C Proof of Theorem 2

The proof is in several steps. At the beginning, I provide the essential tools to construct the equilibrium of the model.[3]

C.1 Step 0: tools for the market makers' updating

Lemma 7. *Let $\mu(t, V)$ be the estimate of the unnormalized density function of the random variable $V = H(\hat{m}_{nT}, nT) A(\hat{m}_{nT}, nT)$ given the stochastic differential equation (1.38) when the insider is informed. Then $\mu(t, V)$ must satisfy the following stochastic differential equation (Zakai equation):*

$$d\mu(t, V) = \frac{\theta\left(t, \tilde{V}\right)}{\sigma_z^2} \mu(t, V) \, dY_t, \qquad \mu(0, V) = f(V),$$

which has a unique solution

$$\mu(t, V) = f(V) \exp\left[\frac{1}{\sigma_z^2}\left(\int_0^t \theta(s, V) \, dY(s) - \frac{1}{2}\int_0^t \theta^2(s, V) \, ds\right)\right].$$

Hence, the value estimate $V(t)$ of $H(\hat{m}_{nT}, nT) A(\hat{m}_{nT}, nT)$ is given by

$$V(t) \equiv \mathbb{E}\left[H(\hat{m}_{nT}, nT) A(\hat{m}_{nT}, nT) | \mathcal{F}_{1,t}\right] = \frac{\int_{\mathrm{V}} V \mu(V, t) \, dV}{\int_{\mathrm{V}} \mu(V, t) \, dV} \tag{A43}$$

where $f(v) = \frac{dF(v)}{dv}$ is the prior probability density function at time 0.

Proof. See Zakai (1969) or Baras (1991). ■

Lemma 8. *The value estimate given by (A43) satisfies the stochastic differential equation*

$$dV(t) = \lambda(t) \left(dY_t - \hat{\theta}(t) \, dt\right), \tag{A44}$$

where

$$\hat{\theta}(t) = \mathbb{E}\left[\theta\left(t, \tilde{V}\right) | \mathcal{F}_{1,t}\right] = \frac{\int_{\mathrm{V}} \theta\left(t, \tilde{V}\right) \mu(V, t) \, dv}{\int_{\mathrm{V}} \mu(V, t) \, dv} \tag{A45}$$

[3] The method of proof is based on Li (2013) that solves the economy with the risk-neutral market makers. He applies the "sequential detection" in the filtering literature.

and

$$\lambda(t) \equiv \frac{\mathbb{E}\left[\theta\left(t,\tilde{V}\right)|\mathcal{F}_{1,t}\right] - V(t)\,\hat{\theta}(t)}{\sigma_z^2}. \tag{A46}$$

In addition,

$$\hat{Y}_{1,t} \equiv Y_t - \int_0^t \hat{\theta}(s)\,ds \tag{A47}$$

is a Brownian Motion with instant variance σ_z^2 under $\mathcal{F}_{1,t}$.

Proof. Applying Ito's Lemma to equation (A43) leads to the above standard filtering results. ■

Through observing the aggregate trading volume Y_t, market makers estimate the probability that Y_t is generated by the the insider has private information or not. This updating problem can be solved as to calculate the likelihood ratio between the two hypotheses, $\delta = 1$ versus $\delta = 0$. Following Li (2013), the logarithm of the likelihood ratio between hypotheses (1.38) and (1.39) is given by

$$\phi(t) \equiv \frac{1}{\sigma_z^2}\left(\int_0^t \left[\hat{\theta}(s) - \theta\left(s,\bar{V}\right)\right] dY(s) - \frac{1}{2}\int_0^t \left[\hat{\theta}^2(s) - \theta^2\left(s,\bar{V}\right)\right] ds\right)$$

where $\hat{\theta}$ is as defined by (A45).

Lemma 9. *The market makers' estimate of the probability that the strategic trader has private information*

$$\pi(t) = \mathbb{E}\left[\delta|\mathcal{F}_t^Y\right] = \frac{\pi_{nT-1}\exp\left[\phi(t)\right]}{1-\pi_{nT-1}+\pi_{nT-1}\exp\left[\phi(t)\right]} \tag{A48}$$

satisfies the following stochastic differential equation:

$$d\pi(t) = \frac{\pi(t)\left[1-\pi(t)\right]}{\sigma_z^2}\left(\hat{\theta}(t) - \theta\left(t,\bar{V}\right)\right) d\hat{Y}(t), \qquad \pi(0) = \pi_{nT-1} \tag{A49}$$

where

$$\hat{Y}(t) = Y_t - \int_0^t \left(\pi(s)\,\hat{\theta}(s) + \left[1-\pi(s)\right]\theta\left(s,\bar{V}\right)\right) ds \tag{A50}$$

is the information process, which is a Brownian motion with instantaneous variance σ_z^2 under the filtration $\mathcal{F}_t^Y$.

Proof. The definition of $\pi(t)$ in equation (A48) is obtained by the Bayes' rule. By Ito's Lemma,

$$\begin{aligned} d\pi(t) &= \pi_\phi d\phi + \frac{1}{2}\pi_{\phi\phi}(d\phi)^2 \\ &= \pi(1-\pi)\,d\phi + \frac{1}{2}\pi(1-\pi)(1-2\pi)(d\phi)^2 \\ &= \frac{\pi(t)\left[1-\pi(t)\right]}{\sigma_z^2}\left(\hat{\theta}(t) - \theta\left(t,\bar{V}\right)\right) d\hat{Y}(t) \end{aligned}$$

where

$$\hat{Y}(t) = Y_t - \int_0^t \left(\pi(s)\hat{\theta}(s) + [1-\pi(s)]\,\theta\left(s, \bar{V}\right)\right) ds.$$

■

Lemma 9 shows that the market makers' probability estimate is governed by a nonlinear stochastic differential equation. Note that when the prior $\pi_{nT-1} = 0$ or $\pi_{nT-1} = 1$, the solution to the belief dynamics (A49) is $\pi(t) \equiv 0$ or $\pi(t) \equiv 1$, respectively.

C.2 Step 1: market makers' updating

First, I show that given the insider's trading strategies when she is informed and not informed, how the market makers estimate the probability that the insider has private information and the price dynamics through nonlinear filtering.

Let $\Pi(t,y)$ be an arbitrary function in $C^{1,2}$ on $[0,1] \times R$ with a close range $[0,1]$. At time $nT-1$, $\log[H(\hat{m}_{nT}, nT)\,A(\hat{m}_{nT}, nT)]$ is normally distributed with mean μ_v and variance $\left(\frac{\beta-\gamma^A}{\beta}\right)^2 \sigma_v^2$. I define $h(y) = \exp\left(\mu_v + \frac{\beta-\gamma^A}{\beta}\lambda y\right)$ and $\bar{V} = e^{\mu_v + \frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2 \sigma_v^2}$. This implies $h^{-1}\left(\bar{V}\right) = \frac{\frac{\beta-\gamma^A}{\beta}\sigma_v^2}{2\lambda}$.

I guess the insider's trading strategy follows

$$\theta(t,y,V) = \frac{h^{-1}(V) - h^{-1}\left(\bar{V}\right) - \Pi(t,y)\left[y - h^{-1}\left(\bar{V}\right)\right]}{nT-t} + \bar{\theta}(t,y), \tag{A51}$$

and

$$\Theta(t,y) = \frac{[1-\Pi(t,y)]\left[y - h^{-1}\left(\bar{V}\right)\right]}{nT-t} + \bar{\theta}(t,y). \tag{A52}$$

The expected trading rate of the insider under market makers' perspective $\mathcal{F}_t^Y$ is

$$\bar{\theta}(t,y) = \frac{\left(\frac{\gamma^A}{\beta}\lambda\sigma_z^2\Pi(t,y) - \Pi_y(t,y)\,\sigma_z^2\right)\cdot E(t,y) + \Pi_y(t,y)\,\sigma_z^2}{\Pi\cdot E(t,y) + 1 - \Pi}, \tag{A53}$$

where I let $E(t,y)$ be defined as[4]

$$E(t,y) = e^{-\frac{\gamma^A}{\beta}\lambda y - \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2 (t-(nT-1))}. \tag{A54}$$

The following Lemma states the market makers' expectation of the insider's order rate and their value estimate of $H(\hat{m}_{nT}, nT)\,A(\hat{m}_{nT}, nT)$, given the insider's order rate $\theta(t,y,V)$ defined in equation (A51).

---

[4] As shown later, Li (2013) is a special case of this economy where $\bar{\theta}(t,y) = 0$ when the market makers are risk-neutral.

Lemma 10. *Let $\hat{Y}_{1,t}$ be a Brownian bridge that satisfies*

$$d\hat{Y}_{1,t} = \left[\theta\left(t, \hat{Y}_{1,t}, V\right) - \Theta\left(t, \hat{Y}_{1,t}\right)\right] dt + dZ_t \tag{A55}$$

$$= \frac{h^{-1}(V) - \hat{Y}_{1,t}}{nT - t} dt + dZ_t \tag{A56}$$

with $\hat{Y}_{1,nT-1} = 0$. If the insider's order rate is $\theta\left(t, \hat{Y}_{1,t}, V\right)$, as defined by (A51), $\Theta\left(t, \hat{Y}_{1,t}\right)$ as defined by (A52), is then the market makers' expected order rate from the insider, conditional on the insider having private information. That is,

$$\hat{\theta}(t) = \mathbb{E}\left[\theta\left(t, \hat{Y}_{1,t}, V\right) | \mathcal{F}_{1,t}\right] = \Theta\left(t, \hat{Y}_{1,t}\right).$$

Furthermore, the expected value of $H(\hat{m}_{nT}, nT) A(\hat{m}_{nT}, nT)$ under $\mathcal{F}_{1,t}$ is

$$V(t) = \mathbb{H}\left(t, \hat{Y}_{1,t}\right),$$

where

$$\mathbb{H}(t, y) = \mathbb{E}\left[h\left(y + Z_{nT} - Z_t\right)\right],$$

where the expectation is taken over the Brownian motion Z.

Proof. See Lemma 6 in Li (2013). ■

From equation (A49), the market makers' estimate of the probability that the insider has private information satisfies

$$d\Pi\left(t, \hat{Y}_{1,t}\right) = \frac{\Pi\left(t, \hat{Y}_{1,t}\right)\left[1 - \Pi\left(t, \hat{Y}_{1,t}\right)\right]}{\sigma_z^2} \frac{\hat{Y}_{1,t} - h^{-1}(\bar{V})}{nT - t} \times \left(d\hat{Y}_{1,t} + \frac{\left[1 - \Pi\left(t, \hat{Y}_{1,t}\right)\right]\left[\hat{Y}_{1,t} - h^{-1}(\bar{V})\right]}{nT - t} dt\right) \tag{A57}$$

with $\Pi\left(nT - 1, \hat{Y}_{1,nT-1}\right) = \pi_{nT-1}$. This equation holds because of Lemma 9 and the following equation:

$$\begin{aligned} dY_t &\overset{(A50),(A61)}{=} d\hat{Y}(t) + \bar{\theta}(t, y) \\ &\overset{(A47)}{=} d\hat{Y}_{1,t} + \Theta(t, y) \\ &\overset{(A52)}{=} d\hat{Y}_{1,t} + \frac{[1 - \Pi(t, y)]\left[y - h^{-1}(\bar{V})\right]}{nT - t} + \bar{\theta}(t, y). \end{aligned} \tag{A58}$$

As shown in Li (2013), $\forall t \in [nT - 1, nT]$, the solution to the stochastic differential equation (A57) is:

$$\Pi\left(t, \hat{Y}_{1,t}\right) = \frac{\pi_{nT-1} \exp\left(\frac{1}{2\sigma_z^2} \frac{\left[\hat{Y}_{1,t} - h^{-1}(\bar{V})\right]^2}{nT - t} + \frac{1}{2}\log(nT - t) - \frac{\left[h^{-1}(\bar{V})\right]^2}{2\sigma_z^2}\right)}{1 - \pi_{nT-1} + \pi_{nT-1} \exp\left(\frac{1}{2\sigma_z^2} \frac{\left[\hat{Y}_{1,t} - h^{-1}(\bar{V})\right]^2}{nT - t} + \frac{1}{2}\log(nT - t) - \frac{\left[h^{-1}(\bar{V})\right]^2}{2\sigma_z^2}\right)}, \tag{A59}$$

which is the market makers' optimal estimate of the probability that the insider has private information, given the insider's order rate $\theta\left(t, \hat{Y}_{1,t}, V\right)$ defined by equation (A51).

When the insider is not better informed, Lemma (10) implies that her trading strategy follows

$$\theta\left(t, y, \bar{V}\right) = -\frac{\Pi\left(t, y\right)\left[y - h^{-1}\left(\bar{V}\right)\right]}{nT - t} + \bar{\theta}\left(t, y\right),$$

which implies

$$\hat{\theta}\left(t\right) - \theta\left(t, y, \bar{V}\right) = \Theta\left(t, y\right) - \theta\left(t, y, \bar{V}\right) = \frac{y - h^{-1}\left(\bar{V}\right)}{nT - t}, \tag{A60}$$

and

$$\Pi\left(t, y\right)\Theta\left(t, y\right) + \left[1 - \Pi\left(t, y\right)\right]\theta\left(t, y, \bar{V}\right) = \bar{\theta}\left(t, y\right). \tag{A61}$$

When the insider has no private information, I can rewrite the dynamics of the probability estimate as

$$\begin{aligned} d\Pi\left(t, \hat{Y}_{1,t}\right) = {} & \frac{\Pi\left(t, \hat{Y}_{1,t}\right)\left[1 - \Pi\left(t, \hat{Y}_{1,t}\right)\right]}{\sigma_z^2}\frac{\hat{Y}_{1,t} - h^{-1}\left(\bar{V}\right)}{nT - t} \\ & \times\left(dZ_t - \frac{\Pi\left(t, \hat{Y}_{1,t}\right)\left[\hat{Y}_{1,t} - h^{-1}\left(\bar{V}\right)\right]}{nT - t}dt\right). \end{aligned} \tag{A62}$$

Therefore, conditional on whether the insider is informed or not, there are two different dynamics of probability estimation, as stated in equations (A57) and (A62).

As stated in Lemma 11, a direct application of Theorem 1 in Li (2013) leads to the same property of probability estimate.

Lemma 11. *Let $\hat{Y}_{1,t}$ be the Brownian bridge as defined by equation (A56) for any $V \in \mathrm{V}$. Suppose that the prior $\pi_{nT-1} \in (0,1)$. Then, the market makers' probability estimate that the insider has private information, $\Pi\left(t, \hat{Y}_{1,t}\right)$, always resides in $(0,1)$ for all $t < nT$. Upon announcements, it converges to 1 or 0 depending on whether the insider has private information or not.*

Since $\log\left[H\left(\hat{m}_{nT}, nT\right)A\left(\hat{m}_{nT}, nT\right)\right]$ is normally distributed with mean μ_v and variance $\left(\frac{\beta-\gamma^A}{\beta}\right)^2\sigma_v^2$ at time $nT-1$, $h\left(y\right) = \exp\left(\mu_v + \frac{\beta-\gamma^A}{\beta}\lambda y\right)$ and $\bar{V} = e^{\mu_v + \frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2\sigma_v^2}$. From Lemma 10, $\forall t \in [nT-1, nT]$, the estimation of $H\left(\hat{m}_{nT}, nT\right)A\left(\hat{m}_{nT}, nT\right)$ conditional on $\delta = 1$ follows

$$\begin{aligned} V\left(t\right) &= \mathbb{E}\left[H\left(\hat{m}_{nT}, nT\right)A\left(\hat{m}_{nT}, nT\right)|\mathcal{F}_{1,t}\right] \\ &= \mathbb{H}\left(t, \hat{Y}_{1,t}\right) = \mathbb{E}\left[h\left(\hat{Y}_{1,t} + Z_{nT} - Z_t\right)\right] \\ &= \exp\left(\mu_V + \frac{\beta-\gamma^A}{\beta}\lambda\hat{Y}_{1,t} + \frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2\sigma_v^2\left(nT - t\right)\right), \end{aligned}$$

while the estimation of $H(\hat{m}_{nT}, nT)\, A(\hat{m}_{nT}, nT)$ conditional on $\delta = 0$ is $\bar{V} = e^{\mu_V + \frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2 \sigma_v^2}$, where $\mu_V = \left(\beta - \gamma^A\right) x_{nT-1} + N(nT)$.

Similarly, the estimation of SDF $H(\hat{m}_{nT}, nT)$ conditional on $\delta = 1$ follows

$$\begin{aligned}\Lambda(t) &= \mathbb{E}\left[H(\hat{m}_{nT}, nT)\,|\mathcal{F}_{1,t}\right] \\ &= \exp\left(\mu_\Lambda - \frac{\gamma^A}{\beta}\lambda \hat{Y}_{1,t} + \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2 (nT - t)\right),\end{aligned}$$

while the estimation of SDF $H(\hat{m}_{nT}, nT)$ conditional on $\delta = 0$ is $\bar{\Lambda} = e^{\mu_\Lambda + \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2}$, where $\mu_\Lambda = -\gamma^A x_{nT-1} + \mathcal{H}(nT)$.

Therefore, this implies the price defined in equation (1.36) depends only on the current adjusted trading flow $\hat{Y}_{1,t}$, which follows

$$\begin{aligned}P\left(t, \hat{Y}_{1,t}\right) &= \frac{\Pi\left(t, \hat{Y}_{1,t}\right) V\left(t, \hat{Y}_{1,t}\right) + \left(1 - \Pi\left(t, \hat{Y}_{1,t}\right)\right) \bar{V}}{\Pi\left(t, \hat{Y}_{1,t}\right) \Lambda\left(t, \hat{Y}_{1,t}\right) + \left(1 - \Pi\left(t, \hat{Y}_{1,t}\right)\right) \bar{\Lambda}} \\ &= \frac{\Pi\left(t, \hat{Y}_{1,t}\right) e^{\mu_V + \frac{\beta-\gamma^A}{\beta}\lambda \hat{Y}_{1,t} + \frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2 \sigma_v^2 (nT-t)} + \left(1 - \Pi\left(t, \hat{Y}_{1,t}\right)\right) e^{\mu_V + \frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2 \sigma_v^2}}{\Pi\left(t, \hat{Y}_{1,t}\right) e^{\mu_\Lambda - \frac{\gamma^A}{\beta}\lambda \hat{Y}_{1,t} + \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2 (nT-t)} + \left(1 - \Pi\left(t, \hat{Y}_{1,t}\right)\right) e^{\mu_\Lambda + \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2}} \\ &= P_{nT-1} \frac{\Pi\left(t, \hat{Y}_{1,t}\right) e^{\frac{\beta-\gamma^A}{\beta}\lambda \hat{Y}_{1,t} - \frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2 \sigma_v^2 (t-(nT-1))} + 1 - \Pi\left(t, \hat{Y}_{1,t}\right)}{\Pi\left(t, \hat{Y}_{1,t}\right) e^{-\frac{\gamma^A}{\beta}\lambda \hat{Y}_{1,t} - \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2 (t-(nT-1))} + 1 - \Pi\left(t, \hat{Y}_{1,t}\right)}, \qquad \text{(A63)}\end{aligned}$$

where $P_{nT-1} = e^{\beta \hat{m}_{nT-1} - \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2 + N(nT) + \frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2 \sigma_v^2}$.

C.3 Step 2: Insider's Optimal Strategy

In this section, I show that if the dynamics of price follows equation (A63), then the optimal trading strategy of the insider is indeed of the form given in equation (A51) through verification proof.

Given the market makers' pricing rule, $P(t) = P\left(t, \hat{Y}_{1,t}\right)$, the insider chooses the order rate to maximize her trading profit. When the insider has private information, for each terminal value $A(\hat{m}_{nT}, nT)$, she maximizes the terminal profit

$$\int_{nT-1}^{nT} \left(A(\hat{m}_{nT}, nT) - P\left(s, \hat{Y}_1(s)\right)\right) \theta_s ds.$$

When the insider is not bettered informed, given no new information coming before announcements, her best estimation of the asset value at $\forall t \in [nT-1, nT]$ is always

$$\begin{aligned}
\bar{v}^* &\equiv \mathbb{E}\left[\frac{H(\hat{m}_{nT}, nT)}{\mathbb{E}\left[H(\hat{m}_{nT}, nT|\mathcal{F}_{nT-1})\right]} A(\hat{m}_{nT}, nT)\,|\mathcal{F}^Y_{nT-1}\right] \\
&= \frac{\mathbb{E}\left[H(\hat{m}_{nT}, nT)\,A(\hat{m}_{nT}, nT)\,|\mathcal{F}^Y_{nT-1}\right]}{\mathbb{E}\left[H(\hat{m}_{nT}, nT)\,|\mathcal{F}^Y_{nT-1}\right]} \equiv \frac{\bar{V}}{\bar{\Lambda}},
\end{aligned} \tag{A64}$$

which is the same as market makers at $t = nT - 1$.

Under Assumption 1, the insider chooses the order rate to maximize the expectation of her terminal profit given the make makers' pricing rule $P(t) = P\left(t, \hat{Y}_{1,t}\right)$:

$$J(t, y; A(\hat{m}_{nT}, nT), \pi_{nT-1}) = \max_{\theta_t \in \mathcal{A}} \mathbb{E}\left[\int_t^{nT} \left(A(\hat{m}_{nT}, nT) - P\left(s, \hat{Y}_s\right)\right)\theta_s ds | \hat{Y}_{1,t} = y, A(\hat{m}_{nT}, nT)\right]$$

subject to

$$d\hat{Y}_t = \left[\theta(t) - \hat{\theta}(t)\right] dt + dZ_t, \tag{A65}$$

where $A(\hat{m}_{nT}, nT) = \bar{v}^*$ when the insider is not informed as shown in equation (A64).

The principle of optimality implies the following Bellman equation

$$\max_{\theta_t \in \mathcal{A}} \left\{ (A(\hat{m}_{nT}, nT) - P(t, y))\,\theta_t + J_t + J_y\left[\theta_t - \hat{\theta}(t)\right] + \frac{1}{2}\sigma_z^2 J_{yy} \right\} = 0 \tag{A66}$$

where the subscripts denote the derivatives. The necessary conditions for having an optimal solution to the Bellman equation (A66) are

$$J_y = P(t, y) - A(\hat{m}_{nT}, nT) \tag{A67}$$

$$J_t + \frac{1}{2}\sigma_z^2 J_{yy} - \hat{\theta}(t)\,J_y = 0. \tag{A68}$$

Under these necessary conditions, a direct application of Li (2013) leads to the following results:

Lemma 12. *Suppose the expected order rate $\hat{\theta}(t) = \Theta\left(t, \hat{Y}_{1,t}\right)$, where $\hat{Y}_{1,t}$ is the adjusted order at t. Let $\omega_t = y$ and suppose that the stochastic differential equation*

$$d\omega_s = dZ_s - \Theta(s, \omega_s)\,ds, \quad \forall nT \geq s \geq t \geq nT - 1$$

has a unique solution, where Z_s is a Brownian motion with instant variance σ_z^2. If there exists a strictly monotone function $g(\cdot)$ such that the pricing rule is

$$P(t, y) = \mathbb{E}\left[g(\omega_{nT})\,|\omega_t = y\right], \tag{A69}$$

then

$$J(t, y; v, \pi_{nT-1}) = \mathbb{E}\left[j(v, \omega_{nT})\,|\omega_t = y\right]$$

is a smooth solution to the Bellman equations (18) and (19), where

$$j(v,y) = \int_{y}^{g^{-1}(v)} [v - g(x)]\, dx \geq 0, \quad \forall (v,y)$$

Lemma 13. *Any continous trading strategy that makes* $\lim_{t \to nT^-} P\left(t, \hat{Y}_{1,t}\right) = A(\hat{m}_{nT}, nT)$ *is optimal, where* $P(t,y)$ *is as defined by equation (A69).*

Equation (A69) implies that $P(t, \omega_t)$ is a martingale under the filtration generated by ω.[5] This implies the price dynamics with respect to $\mathcal{F}_{1,t}$ must satisfy

$$P_t - \Theta(t,y) P_y + \frac{1}{2}\sigma_z^2 P_{yy} = 0.$$

Finally, I am ready to prove that (X_0, X_1, P, Π) is an equilibrium. The insider's trading strategy, $X_{\delta,t}$, satisfies $X_0(t) = \int_{nT-1}^{t} \theta\left(s, \hat{Y}_1(s), \bar{V}\right) ds$ and $X_1(t) = \int_{nT-1}^{t} \theta\left(s, \hat{Y}_1(s), V\right) ds$, where $\hat{Y}_1$ is the solution to the stochastic differential equation (A56). $\Pi(t,y)$ and $\theta(t,y,V)$ are defined by equations (A59) and (A51), respectively.

Proof of Theorem 2. Note that I have established that $\Pi\left(t, \hat{Y}_{1,t}\right)$ is the optimal probability estimate of market makers given the trading strategy in equation (A51). Then, I need to show that the price dynamics in equation (A63) is a legitimate pricing rule. That is,

[1]. The price rule defined above satisfies

$$P_t - \Theta(t,y) P_y + \frac{1}{2}\sigma_z^2 P_{yy} = 0, \tag{A70}$$

[2]. $P\left(nT, \hat{Y}_{1,t}\right)$ is an increasing function of $\hat{Y}_{1,t}$; and

[3]. $\lim_{t \to nT} P\left(t, \hat{Y}_{1,t}\right) = A(\hat{m}_{nT}, nT)$ almost surely.

The first condition can be shown by direct calculation. For convenience, I let

$$P\left(t, \hat{Y}_{1,t}\right) \quad \equiv \quad P_{nT-1} \frac{A\left(t, \hat{Y}_{1,t}\right)}{B\left(t, \hat{Y}_{1,t}\right)}$$

where

$$A\left(t, \hat{Y}_{1,t}\right) = \Pi\left(t, \hat{Y}_{1,t}\right) e^{\frac{\beta - \gamma^A}{\beta}\lambda \hat{Y}_{1,t} - \frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2 \sigma_v^2 (t-(nT-1))} + 1 - \Pi\left(t, \hat{Y}_{1,t}\right),$$

and

$$B\left(t, \hat{Y}_{1,t}\right) = \Pi\left(t, \hat{Y}_{1,t}\right) e^{-\frac{\gamma^A}{\beta}\lambda \hat{Y}_{1,t} - \frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2 \sigma_v^2 (t-(nT-1))} + 1 - \Pi\left(t, \hat{Y}_{1,t}\right).$$

[5] Notice that due to the existence of the SDF, the pricing rule $P(t)$ is no longer a martingale under market makers (unconditional) information set $\mathcal{F}_t^Y$. Though both $V(t)$ and $\Lambda(t)$ are martingales under $\mathcal{F}_t^Y$.

In addition, I let $D(t,y)$ be defined as

$$D\left(t,\hat{Y}_{1,t}\right)=e^{\frac{\beta-\gamma^A}{\beta}\lambda\hat{Y}_{1,t}-\frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2\sigma_v^2(t-(nT-1))}.$$

I also use the definition of $E(t,y)$ from equation (A54).

The first-order conditions and second-order conditions of the price dynamics (A63) are

$$P_t=P_{nT-1}B^{-2}\left\{\left[\left(\Pi_t-\frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2\sigma_v^2\Pi\right)D-\Pi_t\right]B-A\left[\left(\Pi_t-\frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2\sigma_v^2\Pi\right)E-\Pi_t\right]\right\},$$

$$P_y=P_{nT-1}B^{-2}\left\{\left[\left(\Pi_y+\frac{\beta-\gamma^A}{\beta}\lambda\Pi\right)D-\Pi_y\right]B-A\left[\left(\Pi_y-\frac{\gamma^A}{\beta}\lambda\Pi\right)E-\Pi_y\right]\right\},$$

and

$$\begin{aligned}P_{yy}=&-P_{nT-1}B^{-2}\cdot 2B^{-1}\left\{\left[\left(\Pi_y+\frac{\beta-\gamma^A}{\beta}\lambda\Pi\right)D-\Pi_y\right]B-A\left[\left(\Pi_y-\frac{\gamma^A}{\beta}\lambda\Pi\right)E-\Pi_y\right]\right\}\left[\left(\Pi_y-\frac{\gamma^A}{\beta}\lambda\Pi\right)E-\Pi_y\right]+\\&P_{nT-1}B^{-2}\left\{\left[\left(\Pi_{yy}+2\frac{\beta-\gamma^A}{\beta}\lambda\Pi_y+\left(\frac{\beta-\gamma^A}{\beta}\right)^2\lambda^2\Pi\right)D-\Pi_{yy}\right]B-A\left[\left(\Pi_{yy}-2\frac{\gamma^A}{\beta}\lambda\Pi_y+\left(\frac{\gamma^A}{\beta}\right)^2\lambda^2\Pi\right)E-\Pi_{yy}\right]\right\}.\end{aligned}$$

Put these derivatives into the following equation:

$$\begin{aligned}&\frac{B^2}{P_{nT-1}}\left\{P_t-\left(\Theta-\bar{\theta}+\theta\right)P_y+\frac{1}{2}\sigma_z^2P_{yy}\right\}\\
=&\left\{\left[\left(\Pi_t-\frac{1}{2}\left(\frac{\beta-\gamma^A}{\beta}\right)^2\sigma_v^2\Pi\right)D-\Pi_t\right]B-A\left[\left(\Pi_t-\frac{1}{2}\left(\frac{\gamma^A}{\beta}\right)^2\sigma_v^2\Pi\right)E-\Pi_t\right]\right\}-\\
&\left(\Theta-\bar{\theta}+\theta\right)\left\{\left[\left(\Pi_y+\frac{\beta-\gamma^A}{\beta}\lambda\Pi\right)D-\Pi_y\right]B-A\left[\left(\Pi_y-\frac{\gamma^A}{\beta}\lambda\Pi\right)E-\Pi_y\right]\right\}+\\
&\frac{1}{2}\sigma_z^2\left\{\left[\left(\Pi_{yy}+2\frac{\beta-\gamma^A}{\beta}\lambda\Pi_y+\left(\frac{\beta-\gamma^A}{\beta}\right)^2\lambda^2\Pi\right)D-\Pi_{yy}\right]B-A\left[\left(\Pi_{yy}-2\frac{\gamma^A}{\beta}\lambda\Pi_y+\left(\frac{\gamma^A}{\beta}\right)^2\lambda^2\Pi\right)E-\Pi_{yy}\right]\right\}\\
&-B^{-1}\sigma_z^2\left\{\left[\left(\Pi_y+\frac{\beta-\gamma^A}{\beta}\lambda\Pi\right)D-\Pi_y\right]B-A\left[\left(\Pi_y-\frac{\gamma^A}{\beta}\lambda\Pi\right)E-\Pi_y\right]\right\}\left[\left(\Pi_y-\frac{\gamma^A}{\beta}\lambda\Pi\right)E-\Pi_y\right]\\
=&\underbrace{\left[\Pi_t-\left(\Theta-\bar{\theta}\right)\Pi_y+\frac{1}{2}\sigma_z^2\Pi_{yy}\right]}_{\text{term (a)}}\left[B\left(D-1\right)-A\left(E-1\right)\right]+\underbrace{\left[\sigma_z^2\Pi_y-\left(\Theta-\bar{\theta}\right)\Pi\right]}_{\text{term (b)}}\left(\frac{\beta-\gamma^A}{\beta}\lambda DB+\frac{\gamma^A}{\beta}\lambda AE\right)\\
&-\underbrace{\left(\bar{\theta}+B^{-1}\sigma_z^2\left(\left(\Pi_y-\frac{\gamma^A}{\beta}\lambda\Pi\right)E-\Pi_y\right)\right)}_{\text{term (c)}}\left\{\left[\left(\Pi_y+\frac{\beta-\gamma^A}{\beta}\lambda\Pi\right)D-\Pi_y\right]B-A\left[\left(\Pi_y-\frac{\gamma^A}{\beta}\lambda\Pi\right)E-\Pi_y\right]\right\}.\end{aligned}\tag{A71}$$

Next, I will show all of term (a), (b), and (c) are zero under the trading strategy (A51). From equations (A57) and (A58),

$$d\Pi\left(t,\hat{Y}_{1,t}\right)=\frac{\Pi\left(t,\hat{Y}_{1,t}\right)\left[1-\Pi\left(t,\hat{Y}_{1,t}\right)\right]}{\sigma_z^2}\frac{\hat{Y}_{1,t}-h^{-1}\left(\bar{V}\right)}{nT-t}d\hat{Y}\left(t\right),$$

which implies $\Pi\left(t, \hat{Y}_{1,t}\right)$ is a martingale under market makers' information set since $\hat{Y}(t)$ is a Brownian motion under $\mathcal{F}_t^Y$. In addition,

$$\begin{aligned} d\hat{Y}_{1,t} &= dY_t - \Theta\left(t, \hat{Y}_{1,t}\right) dt \\ &= \left[d\hat{Y}(t) + \bar{\theta}\left(t, \hat{Y}_{1,t}\right) dt\right] - \Theta\left(t, \hat{Y}_{1,t}\right) dt \\ &= d\hat{Y}(t) - \left[\Theta\left(t, \hat{Y}_{1,t}\right) - \bar{\theta}\left(t, \hat{Y}_{1,t}\right)\right] dt, \end{aligned} \tag{A72}$$

I have

$$\Pi_t - \left[\Theta(t, y) - \bar{\theta}(t, y)\right] \Pi_y + \frac{1}{2}\sigma_z^2 \Pi_{yy} = 0 \tag{A73}$$

This shows that term (a) in equation (A71) is always zero for any $(t, \hat{Y}_{1,t})$.

Moreover, from equation (A59),

$$\Pi_y = \Pi(1 - \Pi) \frac{y - h^{-1}(\bar{V})}{\sigma_z^2 (nT - t)}. \tag{A74}$$

Combining with equation (A52), I have

$$\sigma_z^2 \Pi_y - \left[\Theta(t, y) - \bar{\theta}(t, y)\right] \Pi = 0, \tag{A75}$$

which implies term (b) in equation (A71) is always zero for any $(t, \hat{Y}_{1,t})$.

The definition of $\bar{\theta}(t, y)$ directly indicates term (c) in equation (A71) is always zero for any $(t, \hat{Y}_{1,t})$. Therefore, the price dynamics satisfy (A70) in the condition [1], which holds for all states of nature.[6]

As $\Pi\left(nT, \hat{Y}_{1,nT}\right) = 1$ when the insider is better informed by Lemma 11, I have

$$P\left(nT, \hat{Y}_{1,t}\right) = P_{nT-1} e^{\lambda \hat{Y}_{1,t} - \frac{1}{2}\frac{\beta - 2\gamma^A}{\beta}\sigma_v^2 (t - (nT-1))},$$

which increases in $\hat{Y}_{1,t}$ since $\lambda > 0$. This verifies the condition [2].

From Lemma 11, when the insider is better informed,

$$\begin{aligned} \lim_{t \to nT} \log P\left(t, \hat{Y}_{1,t}\right) &= \log P_{nT-1} - \frac{1}{2}\frac{\beta - 2\gamma^A}{\beta}\sigma_v^2 + \lambda \lim_{t \to nT} \hat{Y}_{1,t} \\ &= \log P_{nT-1} - \frac{1}{2}\frac{\beta - 2\gamma^A}{\beta}\sigma_v^2 + \beta\left(\hat{m}_{nT} - \hat{m}_{nT-1}\right) \quad \text{a.s.} \\ &= \log A\left(\hat{m}_{nT}, nT\right) \quad \text{a.s.} \end{aligned}$$

The second equality holds since $\hat{Y}_{1,t}$ is a Brownian bridge that converges to $h^{-1}(V)$ almost surely. The third equality comes from the definition of P_{nT-1}. Therefore, the condition [3] also holds.

■

[6] Combining equations (A74) and (A53), I can derive $\bar{\theta}(t, y)$ as in equation (1.43).

Proof of Proposition 3. When the insider is informed, from Lemma 11 and Theorem 2, I show $P\left(t, \hat{Y}_{1,t}\right)$ converges to the true asset value $A\left(\hat{m}_{nT}, nT\right)$ almost surely when $t \rightarrow nT$. Since all the private information is eventually incorporated into the price, it implies there is no uncertainty left just upon announcements:

$$Var\left[\log P_{nT-}|\mathcal{F}^{Y}_{nT-}\right] = 0. \tag{A76}$$

While when the insider is not informed, $P\left(t, \hat{Y}_{1,t}\right)$ converges to the initial price P_{nT-1} almost surely when $t \rightarrow nT$. There is no uncertainty reduction upon announcements since the insider has no information other than what the market makers have at $nT-1$:

$$Var\left[\log P_{nT-}|\mathcal{F}^{Y}_{nT-}\right] = Var\left[\log P_{nT-}|\mathcal{F}^{Y}_{nT-1}\right] = \beta^2 \Delta Q. \tag{A77}$$

Therefore, when η fraction of insider that is informed across these FOMC announcements,

$$\log \mathbb{E}\left[\frac{P_{nT-}}{P_{nT-1}}\right] = \eta\gamma^{A}\beta\Delta Q, \tag{A78}$$

$$\mathbb{E}\left[Var\left[\log P_{nT-}|\mathcal{F}^{Y}_{nT-}\right] - Var\left[\log P_{nT-}|\mathcal{F}^{Y}_{nT-1}\right]\right] = -\eta\beta^2\Delta Q, \tag{A79}$$

where the expectations are taken over all states of natural. ■

Appendix B

Chapter 2 Appendices

A Details of Solution Method

We solve the partial differential equation in (2.40) with a finite difference method that approximates the function $p(c, w)$ on a two-dimensional non-rectangular grid: $c \in \{c_i(w_j)\}_{i=1}^{I^j}$ and $w \in \{w_j(c_i)\}_{j=1}^{J^i}$, where we define $\overline{w}(c_i) = w_{J^i}(c_i)$ and $\overline{c}(w_j) = c_{I^j}(w_j)$. Each set of grid points along j, $w_j(c_i)$, depend on the value of c_i, because of the boundary curve $\{\overline{w}(c_i)\}_{i=1}^{I}$. The set of grid points along i, $c_i(w_j)$ shares the same logic.

We approximate first derivatives of p using both backward and forward differences and second derivatives with central differences. All differences of c and w are calculated respectively over the fixed increments Δ_c and Δ_w. For the approximation of the derivatives at the boundaries, there are three different cases:

1. The boundary conditions of w imply that $p(c, 0) = l_K + l_C c \Rightarrow p(c_i, w_0) \approx l_K + l_C c_i$ and $p_w(c, \overline{w}(c)) = -1 \Rightarrow p(c_i, w_{J^i+1}) \approx p(c_i, w_{J^i}) - \Delta_w$ under a forward difference, where both conditions hold for all i.

2. The boundary conditions of c imply that $p(0, w) = p(f(w), w) - \Phi - (1 + \phi)f(w) \Rightarrow p(c_0, w_j) \approx p(f(w_j), w_j) - \Phi - (1 + \phi)f(w_j)$ and $p_c(\overline{c}(w), w) = 1 \Rightarrow p(c_{I^j+1}, w_j) \approx p(c_{I^j}, w_j) + \Delta_c$ under a forward difference, where both conditions hold for all j.

3. The boundary conditions along the joint upper boundary where $p_{cw}(\overline{c}(w), \overline{w}(c)) = 0$ for all $\overline{c}(w)$ and $\overline{w}(c)$ implies

$$p(c_{I^j+1}, w_{J^i+1}) \approx p(c_{I^j+1}, w_{J^i}) - \Delta_w \approx p(c_{I^j}, w_{J^i}) - \Delta_w + \Delta_c.$$

We describe our computational algorithm below:

1. Guess the value of $p^b\,(c, w)$ on the two-dimensional non-rectangular grid: $c \in \{c_i\}_{i=1}^{I^j}$ and $w \in \{w_j(c_i)\}_{j=1}^{J^i}$ and approximate the derivatives
2. Calculate the investment policy function in (2.41)
3. For each w in $[w_1(c_1), w_{J^1}(c_1)]$, we use bisection to find the refinancing policy $f(w)$ such that $p_c\,(f(w), w) = 1 + \phi$
4. We update the value function through an implicit method that solves the vector $p^{b+1} = (p^{b+1}_{1,1}, \ldots, p^{b+1}_{1,J^1}, p^{b+1}_{2,1}, \ldots, p^{b+1}_{2,J^2}, \ldots, p^{b+1}_{I^J,J^I})'$ with notation $p_{i,j} = p(c_i, w_j)$. It begins with a guess $b = 1$ and proceeds to iterate until convergence $(\max(|p^{b+1} - p^b|) < 10^{-9})$ on the value function

$$p^{b+1}\left[\left(\frac{1}{\Delta} + r - (i - \delta)\right) - \mathbf{Q}\right] = p^b/\Delta + B,$$

where i is calculated from step 3, $\Delta > 0$ is the step size of the iterative method, and $\mathbf{Q}$ is the transition matrix defined by the diffusion processes of the states c and w and the boundaries described above

$$\mathbf{Q} = \begin{bmatrix}
q^{ss}_{1,1} & q^{su}_{1,1} & 0 & \cdots & 0 & q^{us}_{1,1} & q^{uu}_{1,1} & 0 & \cdots & 0 & \cdots & 0 \\
q^{sd}_{1,2} & q^{ss}_{1,2} & q^{su}_{1,2} & \ddots & \vdots & q^{ud}_{1,2} & q^{us}_{1,2} & q^{uu}_{1,2} & \ddots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & \cdots & q^{sd}_{1,J^1} & q^{ss}_{1,J^1} & 0 & \cdots & \cdots & q^{ud}_{1,J^1} & q^{us}_{1,J^1} & \ddots & \vdots \\
q^{ds}_{2,1} & q^{du}_{2,1} & 0 & \cdots & 0 & q^{ss}_{2,1} & q^{su}_{2,1} & 0 & \cdots & 0 & \ddots & \vdots \\
q^{dd}_{2,2} & q^{ds}_{2,2} & q^{du}_{2,2} & \ddots & \vdots & q^{sd}_{2,2} & q^{ss}_{2,2} & q^{su}_{2,2} & \ddots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & \cdots & q^{dd}_{2,J^2} & q^{ds}_{2,J^2} & 0 & \cdots & \cdots & q^{sd}_{2,J^2} & q^{ss}_{2,J^2} & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & q^{sd}_{I^J,J^I} & q^{ss}_{I^J,J^I}
\end{bmatrix}.$$

Adjustments to transition rates along the boundaries are made to $\mathbf{Q}$ for the non-rectangular grid as it is an approximation and ensure that the non-termination-boundary rows of the transition matrix $\mathbf{Q}$ sum to zero. The termination-boundary rows do not sum to zero as they measure the (absorbing) exiting mass of firms. The matrix $\mathbf{Q}$ is the discretized analogy of the infinitesimal generator of (dc_t, dw_t): $\mathcal{A}\vartheta(c, w)$ for some arbitrary function $\vartheta(\cdot)$. The elements of $\mathbf{Q}$ are based on an upwind scheme and defined as

- $q^{ss}_{i,j} = -\max(\mathbb{E}_t[dw], 0)/\Delta_w + \min(\mathbb{E}_t[dw], 0)/\Delta_w - \max(\mathbb{E}_t[dc], 0)/\Delta_c + \min(\mathbb{E}_t[dc], 0)/\Delta_c - \mathbb{E}_t[dw^2]/\Delta_w^2 - \mathbb{E}_t[dc^2]/\Delta_c^2$
- $q^{su}_{i,j} = \max(\mathbb{E}_t[dw], 0)/\Delta_w + \mathbb{E}_t[dw^2]/(2\Delta_w^2)$
- $q^{sd}_{i,j} = -\min(\mathbb{E}_t[dw], 0)/\Delta_w + \mathbb{E}_t[dw^2]/(2\Delta_w^2)$
- $q^{us}_{i,j} = \max(\mathbb{E}_t[dc], 0)/\Delta_c + \mathbb{E}_t[dc^2]/(2\Delta_c^2)$
- $q^{ds}_{i,j} = -\min(\mathbb{E}_t[dc], 0)/\Delta_c + \mathbb{E}_t[dc^2]/(2\Delta_c^2)$
- $q^{uu}_{i,j} = -q^{du}_{ij} = -q^{ud}_{ij} = q^{dd}_{ij} = \mathbb{E}_t[dwdc]/(4\Delta_c\Delta_w)$,

where the conditional moments of state variables are $\mathbb{E}_t[dw] = (\gamma - (i - \delta))w$, $\mathbb{E}_t[dc] = ((1 - \tau_Y)\mu - g(i) + \delta\tau_Y + [r(1 - \tau_C) - (i - \delta)]c)$, $\mathbb{E}_t[dw^2] = (\sigma\lambda(c)/\mu)^2$, $\mathbb{E}_t[dc^2] = (\sigma(1 - \tau_Y))^2$, and $\mathbb{E}_t[dcdw] = \sigma^2(1 - \tau_Y)\lambda(c)/\mu$.

Lastly, the vector of constants B required by the boundaries takes the form

$$B = \begin{bmatrix} (q^{dd}_{1,1} + q^{ud}_{1,1} + q^{sd}_{1,1}) \times (l_K + l_C c_1) \\ \vdots \\ (q^{su}_{1,J^1}) \times (-\Delta_w) \\ (q^{dd}_{2,1} + q^{ud}_{2,1} + q^{sd}_{2,1}) \times (l_K + l_C c_2) \\ \vdots \\ (q^{su}_{2,J^2}) \times (-\Delta_w) \\ \vdots \\ \vdots \\ (q^{su}_{I^J,J^I}) \times (-\Delta_w) \end{bmatrix} + \begin{bmatrix} (q^{ds}_{1,1} + q^{dd}_{1,1} + q^{du}_{1,1}) \times p_{1,1} \\ (q^{ds}_{1,2} + q^{dd}_{1,2} + q^{du}_{1,2}) \times p_{1,2} \\ \vdots \\ (q^{ds}_{1,J^1} + q^{dd}_{1,J^1} + q^{du}_{1,J^1}) \times p_{1,J^1} \\ \vdots \\ (q^{us}_{I^1,1}) \times (\Delta_c) \\ \vdots \\ \vdots \\ (q^{us}_{I^J,J^I}) \times (\Delta_c) \end{bmatrix},$$

and intuitively captures the rates of cash outflows from payments to managers, $-\Delta_w$, and inflows to investors from payouts, Δ_c, liquidation, $l_K + l_C c$, and refinancing, $p_{1,j}$ for $j = 1, \ldots, J^1$, where we use the equilibrium value matching condition for refinancing, (2.30).

Stationary Distribution

The stationary distribution, $h(c,w)$, is calculated by solving, $h(c,w) = -(\mathbf{Q}^T)^{-1}\psi$, where ψ is the entry vector. The rows of ψ that are non-zero are determined by the assumed shape of the entry distribution that isolates c and the assumption on how agents' initial continuation utility w is determined through bargaining power. The normalization of $h(c,w)$ to one implies that the entry rate equals $m = -\sum_i \mathbf{Q}^T h(c,w)\Delta_w \Delta_c$.

In the model with no agency friction there is no exit as firms always refinance. Here we compute the stationary distribution $h(c)$ by solving an eigenvalue problem of the adjoint of the transition matrix $\mathbf{Q}$: $\mathbf{Q}^T h(c) = \mathbf{0}$.

B Proof of Tradeoff Along the Joint Upper Boundary

We first prove a lemma on investment being constant along the joint upper boundary before turning to prove the proposition. In what follows we ignore dependence of boundaries on states for brevity: that is, $\overline{c} = \overline{c}(w)$ and $\overline{w} = \overline{w}(c)$.

Lemma. *Investment is constant along the joint upper boundary.*

Proof. Denote $(\overline{c}, \overline{w})$ and $(\overline{c} + dc, \overline{w} + dw)$ as two points along the joint upper boundary. From the first-order condition in (2.41) we have

$$g'(i(c,w)) = \frac{p(c,w) - p_w(c,w)w}{p_c(c,w)} - c,$$

where optimality implies that the investment rates at these two points are

$$g'(i(\overline{c},\overline{w})) = p(\overline{c},\overline{w}) + \overline{w} - \overline{c}, \text{ and} \tag{A1}$$

$$g'(i(\overline{c}+dc,\overline{w}+dw)) = p(\overline{c}+dc,\overline{w}+dw) + (\overline{w}+dw) - (\overline{c}+dc). \tag{A2}$$

Next, from the continuity of the value function along the boundary we know

$$\frac{p(\overline{c},\overline{w}+dw) - p(\overline{c},\overline{w})}{dw} = -1 \qquad \text{and} \qquad \frac{p(\overline{c}+dc,\overline{w}+dw) - p(\overline{c},\overline{w}+dw)}{dc} = 1,$$

which we can rearrange to yield

$$p(\overline{c},\overline{w}) = p(\overline{c},\overline{w}+dw) + dw = p(\overline{c}+dc,\overline{w}+dw) - dc + dw \tag{A3}$$

and then adding $\overline{w} - \overline{c}$ to both sides of (A3) gives

$$p(\overline{c},\overline{w}) + \overline{w} - \overline{c} = p(\overline{c}+dc,\overline{w}+dw) + (\overline{w}+dw) - (\overline{c}+dc). \tag{A4}$$

Finally, the optimal investment rates along the boundary ((A1) and (A2)) and (A4) imply

$$i(\bar{c}, \overline{w}) = i(\bar{c} + dc, \overline{w} + dw).$$

Therefore, the investment rate is constant along the joint upper boundary. ■

Now we turn to proving the proposition, which we restate here for convenience.

Proposition (Tradeoff Along the Joint Upper Boundary)**.** *Consider a marginal change along the joint upper boundary from* $(\bar{c}(w), \overline{w}(c))$ *to* $(\bar{c}(w) + dc, \overline{w}(c) + dw)$*, then the rate of change across this boundary is equal to*

$$\frac{dw}{dc} = -\frac{r\tau_C}{\gamma - r} < 0$$

Proof. The partial differential equation in (2.40) is

$$\begin{aligned} rp(c,w) = \max_i p(c,w)(i-\delta) + p_c(c,w)\left((1-\tau_Y)\mu - g(i) + \delta\tau_Y + [r(1-\tau_C) - (i-\delta)]c\right) \\ + p_w(c,w)(\gamma - (i-\delta))w + \frac{1}{2}p_{cc}(c,w)(\sigma(1-\tau_Y))^2 \\ + \frac{1}{2}p_{ww}(c,w)\left(\frac{\sigma}{\mu}\lambda(c)\right)^2 + p_{cw}(c,w)\frac{\sigma^2(1-\tau_Y)}{\mu}\lambda(c). \end{aligned} \tag{A5}$$

Using our two points, $(\bar{c}, \overline{w})$ and $(\bar{c} + dc, \overline{w} + dw)$, we can reduce (A5) to

$$\begin{aligned} p(\bar{c}, \overline{w})[r - (i-\delta)] = -(\gamma - (i-\delta))\overline{w} \\ + (1-\tau_Y)\mu - g(i) + \delta\tau_Y + [r(1-\tau_C) - (i-\delta)]\bar{c} \end{aligned} \tag{A6}$$

and

$$\begin{aligned} p(\bar{c} + dc, \overline{w} + dw)[r - (i-\delta)] = -(\gamma - (i-\delta))(\overline{w} + dw) \\ + (1-\tau_Y)\mu - g(i) + \delta\tau_Y + [r(1-\tau_C) - (i-\delta)](\bar{c} + dc), \end{aligned} \tag{A7}$$

where we use the Lemma to simplify investment $i \equiv i(\bar{c}, \overline{w}) = i(\bar{c} + dc, \overline{w} + dw)$. Therefore, when subtracting (A6) from (A7) we are left with

$$[r - (i-\delta)](p(\bar{c} + dc, \overline{w} + dw) - p(\bar{c}, \overline{w})) = [r(1-\tau_C) - (i-\delta)]dc - (\gamma - (i-\delta))dw. \tag{A8}$$

We can then place (A3) into (A8) to find

$$[r - (i-\delta)](dc - dw) = [r(1-\tau_C) - (i-\delta)]dc - (\gamma - (i-\delta))dw,$$

which reduces to (2.42) in the text. ■

C Derivations of Distortions

We derive the conditional expectation for the agency distortion and use but do not report for brevity a similar derivation for the financial distortion (2.49). We ignore the dependence of $\mu_c(c, w)$ on state variables in what follows.

$$\begin{aligned}
&\mathbb{E}[\Delta c \mathbf{1}\{\bar{c} < c + \Delta c < \mathcal{F}^{-1}(w)\}|c, w] \\
&= \int_{-\infty}^{\infty} \Delta c \frac{1}{\sqrt{2\pi}\sigma_c} e^{-\frac{1}{2}\left(\frac{\Delta c - \mu_c}{\sigma_c}\right)^2} \mathbf{1}\left\{\bar{c}(w) - c < \Delta c < \mathcal{F}^{-1}(w) - c\right\} d(\Delta c) \qquad \text{(A9)}
\end{aligned}$$

We then use the change of variable $\Delta x = (\Delta c - \mu_c)/\sigma_c$ to change (A9) to

$$\begin{aligned}
&\int_{-\infty}^{\infty} (\sigma_c \Delta x + \mu_c) \frac{1}{\sqrt{2\pi}\sigma_c} e^{-\frac{1}{2}(\Delta x)^2} \mathbf{1}\left\{\bar{c}(w) - c < \sigma_c \Delta x + \mu_c < \mathcal{F}^{-1}(w) - c\right\} d\left(\sigma_c \Delta x + \mu_c\right) \\
&= \int_{-\infty}^{\infty} (\sigma_c \Delta x + \mu_c) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\Delta x)^2} \mathbf{1}\left\{\frac{\bar{c}(w) - c - \mu_c}{\sigma_c} < \Delta x < \frac{\mathcal{F}^{-1}(w) - c - \mu_c}{\sigma_c}\right\} d\left(\Delta x\right) \\
&= \int_{\frac{\bar{c}(w) - c - \mu_c}{\sigma_c}}^{\frac{\mathcal{F}^{-1}(w) - c - \mu_c}{\sigma_c}} (\sigma_c \Delta x + \mu_c) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\Delta x)^2} d\left(\Delta x\right) \\
&= \frac{\sigma_c}{\sqrt{2\pi}} \int_{\frac{\bar{c}(w) - c - \mu_c}{\sigma_c}}^{\frac{\mathcal{F}^{-1}(w) - c - \mu_c}{\sigma_c}} \Delta x e^{-\frac{1}{2}(\Delta x)^2} d\left(\Delta x\right) + \mu_c \int_{\frac{\bar{c}(w) - c - \mu_c}{\sigma_c}}^{\frac{\mathcal{F}^{-1}(w) - c - \mu_c}{\sigma_c}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\Delta x)^2} d\left(\Delta x\right). \qquad \text{(A10)}
\end{aligned}$$

Since under the change $u = -\frac{1}{2}x^2$ we have $\int x e^{-\frac{1}{2}x^2} dx = -\int e^u du = -e^u + k = -e^{-\frac{1}{2}x^2} + k$, where k is a constant of integration, equation (A10) equals

$$\begin{aligned}
&- \frac{\sigma_c}{\sqrt{2\pi}} e^{-\frac{1}{2}(\Delta x)^2} \Big|_{\frac{\bar{c}(w) - c - \mu_c}{\sigma_c}}^{\frac{\mathcal{F}^{-1}(w) - c - \mu_c}{\sigma_c}} + \mu_c \int_{\frac{\bar{c}(w) - c - \mu_c}{\sigma_c}}^{\frac{\mathcal{F}^{-1}(w) - c - \mu_c}{\sigma_c}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\Delta x)^2} d\left(\Delta x\right) \\
&= \frac{\sigma_c}{\sqrt{2\pi}} \left[e^{-\frac{1}{2}\left(\frac{\bar{c}(w) - c - \mu_c}{\sigma_c}\right)^2} - e^{-\frac{1}{2}\left(\frac{\mathcal{F}^{-1}(w) - c - \mu_c}{\sigma_c}\right)^2} \right] + \mu_c \left[\mathcal{N}\left(\frac{\mathcal{F}^{-1}(w) - c - \mu_c}{\sigma_c}\right) - \mathcal{N}\left(\frac{\bar{c}(w) - c - \mu_c}{\sigma_c}\right) \right].
\end{aligned}$$

D Verification of HJB Optimality and Full Effort Condition

Define the gain process $\{\mathcal{G}\}$ under any incentive-compatible contract $\mathcal{C} = (I, F, D, U, \tau)$ for any $t \leq \tau$ as

$$\mathcal{G}_t(\mathcal{C}) = \int_0^t e^{-rs}(dD_s - dF_s - dX_s - dU_s) + e^{-rt} P(K_t, C_t, W_t),$$

where K_t, C_t, and W_t evolve as in (2.1), (2.4), and (2.8). Homogeneity and Ito's lemma imply

$$e^{rt}d\mathcal{G}_t = K_t \left\{ \begin{array}{c} \left[\begin{array}{c} -rp + p(i_t - \delta) + p_c\left((1-\tau_Y)\mu - g(i_t) + \delta\tau_Y + [r(1-\tau_C) - (i_t - \delta)]c_t\right) \\ +p_w(\gamma - (i_t - \delta))w_t + \frac{1}{2}p_{cc}(\sigma(1-\tau_Y))^2 + \frac{1}{2}p_{ww}\left(\beta_t(1-\tau_Y)\sigma\right)^2 + p_{cw}\beta_t(1-\tau_Y)^2\sigma^2 \end{array} \right] dt \\ + \left[(1-p_c)\times(dD_t - dF_t) - dX_t - (1+p_w)dU_t\right]/K_t + (p_c + \beta_t p_w)\sigma(1-\tau_Y)dZ_t \end{array} \right\},$$

where $p(\cdot)$'s dependence on states (c_t, w_t) and $\mathcal{G}(\cdot)$'s dependence on a contract $\mathcal{C}$ have been omitted for conciseness from this point on.

Under the optimal investment i_t^*, and incentive policies $\beta_t^* = \lambda(c_t)/\left((1-\tau_Y)\mu\right)$ the top two lines in the square brackets are the optimized PDE in (2.40) and therefore zero. For models in which the only state variable is agents' continuation utility this nonpositivity condition follows from the concavity of $p(w)$. In this more general case, we show and verify numerically that for any other incentive-compatible policy both p_{ww} and the sum $\beta p_{ww}/2 + p_{cw}$, under the policy with the smallest β, are nonpositive.[1] Panels A and B of Figure A-1 depict p_{ww} and the sum under β_t^* and the calibration in Table 2.1.

The figure broadly shows that these terms are negative across the entire state space. However, in the upper and lower 10 percent of the distributions, indicated by the dashed lines, there are instances of these terms being slightly positive, especially in the lower portion of w's distribution. Solving this model is difficult, particularly since the exact shape of the boundary of the state space is not well understood and functional forms must be used to approximate it. Accordingly, these instances coincide with regions where super contact conditions diverge further from zero (see Panels C and D of Figure 2.3) and are therefore likely due in part to numerical error and also coincide with regions where the stationary density does not place a large mass (see Panel B of Figure 2.4). Because of this latter fact, perturbing the boundary to minimize further this error has little impact on the quantitative predictions of our model. In our defense, there is no other evidence of non-optimality in the solution and we place no restrictions on agents' ability to process information (for example, Krusell and Smith (1998)).

Next, the term capturing the optimality of the continuation payment policy, $-(1+p_w)dU_t$, is nonpositive since $p_w \geq -1$ but equals zero under the optimal contract. The term that captures the optimality of net funds dispensed, $(1-p_c)(dD_t - dF_t)$, is also nonpositive since (i) $p_c \geq 1$ and (ii) $dD_t \geq dF_t$ because either (ii.a) cash holdings are sufficient for payouts, $dD > 0$ and $dF = 0$; or (ii.b) if current cash holdings are insufficient to finance payouts the firm will raise funds externally for payouts, in which case $dD \geq dF$ since cash holdings are positive. This term is also zero under the optimal contract. Lastly, the issuance cost, $-dX_t$, is nonpositive but

[1] This can be seen by simplifying the second-order terms that depend on the incentive coefficient β, $p_{ww}\beta^2(1-\tau_Y)^2\sigma^2/2 + p_{cw}\beta(1-\tau_Y)^2\sigma^2$, to be less than or equal to zero.

equals zero under the optimal contract.

Therefore, for the auxiliary gain process we have

$$d\mathcal{G}_t = \mu_{\mathcal{G}}(t)dt + e^{-rt}K_t(p_c + \beta_t p_w)\sigma(1-\tau_Y)dZ_t,$$

where $\mu_{\mathcal{G}}(t) \leq 0$. Let $\varphi_t \equiv e^{-rt}K_t(p_c + \beta_t p_w)\sigma(1-\tau_Y)$. We impose the usual regularity conditions to ensure that $\mathbb{E}\left[\int_0^T \varphi_t dZ_t\right] = 0$ for all $T > 0$. This implies that $\{\mathcal{G}\}$ is a supermartingale.

We can now evaluate the principal's payoff for an arbitrary incentive compatible contract. Recall that $P(K_\tau, C_\tau, W_\tau) = l_K K_\tau + l_C C_\tau$. Given any $t < \infty$,

$$\begin{aligned}
&\mathbb{E}\left[\int_0^\tau e^{-rs}(dD_s - dF_s - dX_s - dU_s) + e^{-r\tau}\left(l_K K_\tau + l_C C_\tau\right)\right] \\
&= \mathbb{E}\left[\mathcal{G}_{t\wedge\tau} + 1_{t\leqslant\tau}\left(\int_t^\tau e^{-rs}(dD_s - dF_s - dX_s - dU_s) + e^{-r\tau}\left(l_K K_\tau + l_C C_\tau\right) - e^{-rt}P(K_t, C_t, W_t)\right)\right] \\
&= \mathbb{E}\left[\mathcal{G}_{t\wedge\tau}\right] + \\
&e^{-rt}\mathbb{E}\left[1_{t\leqslant\tau}\left(\mathbb{E}_t\left[\int_t^\tau e^{-r(s-t)}(dD_s - dF_s - dX_s - dU_s) + e^{-r(\tau-t)}\left(l_K K_\tau + l_C C_\tau\right)\right] - P(K_t, C_t, W_t)\right)\right] \\
&\leq \mathcal{G}_0 + \left(q^{FB} - (l_K + l_C \times c)\right)\mathbb{E}[e^{-rt}K_t].
\end{aligned}$$

The first term of the inequality follows from the nonpositive drift of $d\mathcal{G}_t$ and the martingale property of $\int_0^{t\wedge\tau} \varphi_s dZ_s$. The second term follows from

$$\mathbb{E}_t\left[\int_t^\tau e^{-r(s-t)}(dD_s - dF_s - dX_s - dU_s) + e^{-r(\tau-t)}\left(l_K K_\tau + l_C C_\tau\right)\right] \leq q^{FB}K_t - w_t K_t,$$

which is the first-best result and

$$q^{FB}K_t - w_t K_t - P(K_t, C_t, W_t) \leq (q^{FB} - p(c,0))K_t = \left(q^{FB} - (l_K + l_C \times c)\right)K_t.$$

as $w + p(c,w)$ is increasing in w since $p_w(c,w) \geq -1$.

We impose the standard transversality conditions $\lim_{T\to\infty}\mathbb{E}[e^{-rT}K_T] = 0$ and $\lim_{T\to\infty}\mathbb{E}[e^{-rT}C_T] = 0$. Therefore letting $t \to \infty$

$$\mathbb{E}\left[\int_0^\tau e^{-rs}(dD_s - dF_s - dX_s - dU_s) + e^{-r\tau}\left(l_K K_\tau + l_C C_\tau\right)\right] \leq \mathcal{G}_0. \tag{A11}$$

for all incentive-compatible contracts. On the other hand, under the optimal contract $\mathcal{C}^*$, principal's payoff $\mathcal{G}(\mathcal{C}^*)$ achieves $\mathcal{G}_0$ because the above weak inequality holds with equality when $t \to \infty$.

Full Effort Condition

Finally, we require $\Lambda(K_t, C_t)$ to be sufficiently small to ensure the optimality of $e_t = 1$ all the time. When managers shirk ($e_t = 0$) they enjoy private benefits at rate $\Lambda(K_t, C_t)dt$. Cash holdings would then evolve as

$$dC_t = (1-\tau_Y)\sigma K_t dZ_t - I_t dt - G(I_t, K_t)dt + \tau_Y \delta K_t dt + r(1-\tau_C)C_t dt + dF_t - dD_t.$$

When they shirk their payoff would not depend on cash flow realizations, so their continuation payoff would change according to

$$dW_t = \gamma W_t dt - dU_t - \Lambda(K_t, C_t)dt.$$

For this not to be the case and for effort ($e_t = 1$) to remain optimal, it must be that investors' payoff rate from allowing agents to shirk be lower than under the optimal contract and equivalently that investors' optimal gain process remain a supermartingale with respect to this shirking policy:

$$\begin{aligned} rp \geq\ & p(i-\delta) + p_c\left(-g(i) + \delta\tau_Y + [r(1-\tau_C) - (i-\delta)]c\right) \\ & + p_w(\gamma - \lambda(c) - (i-\delta))w + \frac{1}{2}p_{cc}(\sigma(1-\tau_Y))^2, \text{ for all } c \text{ and } w. \end{aligned}$$

We confirm this optimality of full effort numerically and depict the result in Panels C and D of Figure A-1. In more general situations where the inequality binds, then a more complicated contract than the one described in this paper would need to be considered. Zhu (2013) considers this extended contracting environment in the context of the DeMarzo and Sannikov (2006) model.

E Alternative Setup Where Managers are Paid Out of Cash

In this appendix we derive the necessary equations required for the optimality of the equilibrium under the assumption that managers' incremental payments dU subtract from incremental cash holdings dC. We streamline its presentation and elaborate after on the key differences from our benchmark setup.

Under this alternative setup, cash holdings possess the law of motion

$$dC_t = dY_t + \tau_Y \delta K_t dt + r(1-\tau_C)C_t dt + dF_t - dD_t - dU_t$$

and investors maximize

$$P(K_0, C_0, W_0) = \max_{\mathcal{C}} \mathbb{E}\left[\int_0^\tau e^{-rt}(dD_t - dF_t - dX_t) + e^{-r\tau}\left(l_K K_\tau + l_C C_\tau\right)\right]$$

subject to similar conditions as in (2.6). The laws of motion for K and W are identical to the benchmark's. The scaled HJB equation then becomes under $dU = dF = dD = dX = 0$ within the boundaries

$$rp(c,w) = \max_i p(c,w)(i-\delta) + p_c(c,w)((1-\tau_Y)\mu - g(i) + [r(1-\tau_C)-(i-\delta)]c) + p_w(c,w)(\gamma - i + \delta)w$$
$$+ \frac{1}{2}p_{cc}(\sigma(1-\tau_Y))^2 + \frac{1}{2}p_{ww}\left(\frac{\sigma}{\mu}\lambda(c)\right)^2 + p_{cw}(c,w)\frac{\sigma^2(1-\tau_Y)}{\mu}\lambda(c),$$

which is identical to the benchmark's. We enumerate the boundary conditions below.

1. The termination boundary is

$$p(c,0) = l_K + l_C \times c \text{ for all } c.$$

2. For each c, there is a compensation level $\overline{w}(c)$ at which it is optimal to pay managers in current payments,

$$p(c,w) = p(c-(w-\overline{w}(c)),\overline{w}(c)) - (w-\overline{w}(c)) \text{ for } w \geq \overline{w}(c)$$
$$\Rightarrow \frac{p(c,w)-p(c,\overline{w}(c))}{w-\overline{w}(c)} = \frac{p(c-(w-\overline{w}(c)),\overline{w}(c))-p(c,\overline{w}(c))}{w-\overline{w}(c)} - 1$$
$$\Rightarrow p_w(c,\overline{w}(c)) = -p_c(c,\overline{w}(c)) - 1, \text{ as } w \to \overline{w}(c), \tag{A12}$$

and which requires the condition $p_{ww}(c,\overline{w}(c)) = 0$ for each c.

3. When cash holdings reach zero, the firm refinances with an equity issue of size fK. The refinancing decision is determined in part by where w lies relative to $\overline{w}(f)$, managers' payment boundary for each level of post-refinancing cash holdings.

   - If $w \in [0,\overline{w}(f)]$, then for each w

   $$p(0,w) = p(f,w) - \Phi - (1+\phi)f, \tag{A13}$$

   along with the condition $p_c(f,w) = 1+\phi$ as before.

   - If $w > \overline{w}(f)$, then the firm will refinance and pay management current payments, so for each w

   $$p(0,w) = p(f,\overline{w}(f)) - \Phi - (1+\phi)f - (w-\overline{w}(f)). \tag{A14}$$

   Differentiating this equation with respect to c gives

   $$p_c(f,\overline{w}(f)) + p_w(f,\overline{w}(f))\left.\frac{\partial\overline{w}(c)}{\partial c}\right|_{c=f} + \left.\frac{\partial\overline{w}(c)}{\partial c}\right|_{c=f} = 1+\phi. \tag{A15}$$

Because this first-order condition depends only on the marginal value of cash and does not depend on w, we can use continuity of w approaching $\overline{w}(c)$ in (A12) to write

$$p_w(f,\overline{w}(f)) = p_w(c,\overline{w}(c))|_{c=f} = -1 - p_c(c,\overline{w}(c))|_{c=f} = -1 - p_c(f,\overline{w}(f))$$

and combining with (A15) we get

$$p_c(f,\overline{w}(f))\left(1 - \left.\frac{\partial\overline{w}(c)}{\partial c}\right|_{c=f}\right) = 1+\phi. \tag{A16}$$

4. When cash holdings get large we have as before

$$p_c(\overline{c}(w),w) = 1 \text{ and } p_{cc}(\overline{c}(w),w) = 0 \text{ for each } w.$$

5. And finally the mixed derivatives along the boundaries require that

$$\begin{aligned} p_{cw}(\overline{c}(w),w) &= 0 \text{ for each } w, \\ p_{cw}(c,\overline{w}(c)) &= -p_{cc}(c,\overline{w}(c)) \text{ for each } c, \text{ and} \\ p_{cw}(\overline{c}(w),\overline{w}(c)) &= -p_{cc}(\overline{c}(w),\overline{w}(c)) = 0 \text{ for every } c \text{ and } w. \end{aligned}$$

We now discuss the salient differences between the alternative setup and our benchmark. From bullet 5., the mixed derivative at the payment boundary $p_{cw}(c,\overline{w}(c)) = -p_{cc}(c,\overline{w}(c))$ no longer equals zero because it now needs to account for the reduction in cash holdings.

From 2., $p_w(c,w) + p_c(c,w) \geq -1$ rather than simply $p_w(c,w) \geq -1$, showing the intuitive change that the bound of the marginal cost of compensation now depends on the marginal cost of cash. Along the payout boundary $\overline{c}(w)$ the inequality collapses to $p_w(\overline{c}(w),w) \geq -2$: that is, it now costs investors at most two dollars to raise managers' continuation utility marginally—the reduction in cash holdings costs investors one dollar ($p_c(\overline{c}(w),w) = 1$) and raising management's continuation utility costs, at most, another dollar. Because (A12) implies that $p_c(c,\overline{w}(c))$ decreases in c, the slope along the payment boundary, $\partial\overline{w}(c)/\partial c$, should be negative.

At last from 3., refinancing can now follow two decision rules depending on the location of w relative to $\overline{w}(f)$. This hypothetical decision is depicted in Figure A-2. The refinancing decision $f(w)$ is traced out with the dashed line. Refinancing decisions satisfy (A13) as in our benchmark and is depicted by the arrow from $p(0,w)$ to $p(f,w)$, where f is determined by $p_c(f,w) = 1+\phi$. At some w, however, it may be optimal to refinance and concurrently pay managers. This decision is depicted by the top two arrows, first moving from $p(0,w)$ to $p(f,w)$, reflecting the post-refinancing gain in value, and then downwards from w to $\overline{w}(f)$, reflecting the transfer from investors to managers of size $w - \overline{w}(f)$. Along this two-part arrow, the refinancing decision is determined by (A15). The firm would decide by choosing the maximum of (A13) and (A14).

Of course, this is not the only possibility for refinancing. A different equilibrium could be imposed by simply requiring (A13) to hold at all points in the alternative setup. Under this policy, refinancing and current payments to managers would never co-occur, as it is in our benchmark. However, in this alternative setup this decision may not be optimal on behalf of investors as they might prefer to refinance more and pay managers, whereas in our benchmark it is optimal.

To sum up, there are several reasons that we prefer our benchmark to this alternative setup. First, the HJB equations are identical. Second, the alternative setup introduces multiple choices in the refinancing decision and it is unclear a priori how to select the correct choices. Third, the refinancing region satisfying (A14) is unlikely to matter quantitatively, since the stationary distribution is likely to put effectively zero mass on low cash, high manager stake firms because of the optimal contract placing a perfect correlation structure across dc_t and dw_t. Altogether it introduces more complexity into an already challenging setup and is unlikely to make a quantitative impact on our results. That said, we want to acknowledge this shortcoming of our model and given this discussion, our model is likely to approximate mature firms best and not small startups. A model more suited to describing the economics of startups would be in Hartman-Glaser, Mayer, and Milbradt (2019).

Last but not least, the formula describing the tradeoff along the joint upper boundary is little changed in the alternative setup. The full derivation closely follows that for the benchmark model and we therefore do not present it here. To summarize the differences, recall that the bound on compensation is now $p_w(c,w)+p_c(c,w) \geq -1$ and can reach a minimum of $p_w(\overline{c}(w),\overline{w}(c)) = -2$. Using this condition rather than $p_w(\overline{c}(w),\overline{w}(c)) = -1$ changes the slope along the joint upper boundary to be

$$\frac{dw}{dc} = -\frac{r\tau_C}{2(\gamma - r)} < 0.$$

As we discuss in Section 2.4 our model is better suited to measuring relative and not absolute distortions. Therefore, a change to this alternative setup will not impact our steady state analysis of the changes in relative distortions.

F Data Appendix

We use all industrial, standard format, consolidated accounts of firms in Compustat. We exclude firms without a NAICS code and in the utilities (22), financial (52-53), other (91), and public (92) industries. As is standard in the literature, we remove firms with missing or non-positive book assets (*at*) or sales (*sale*) and those with net property, plant, and equipment (*ppent*) of less than five million dollars. Our data sample starts in 1993, when compensation data from Execucomp becomes virtually complete, and ends in 2017. Following Gillan, Hartzell, Koch, and Starks (2018) and McKeon (2015), we exclude observations where *salary* is available yet the item *tdc*1 is missing to minimize backfilling bias and define refinancing as common stock issuance greater than 5 percent of book assets. Because in the model state variables are defined over K and in the data over assets $(C + K)$, we subtract cash from assets in the data to make variables comparable and define *net assets* as book assets less cash, $at - che$.

Cash Holdings	= cash (che) / net assets
Compensation	= (salary + bonus + LTIP + equity rewards) (tdc1(t)) / net assets (t-1)
Payout	= 1 if dvc > 0 or repurchases > 0; 0 otherwise
Free Cash Flow	= (EBITDA (ebitda(t)) - physical investment (capx(t))) / net assets (t-1)
Investment	= (physical investment (capx(t)) / net assets (t-1)
Preferred Issuance	= Use max(pstkr(t) - pstkr(t-1),0), max(pstk(t) - pstk(t-1),0), or zero, in decreasing order of preference
Refinancing	= 1 if sale of common stock (sstk less preferred issuance) / assets (at) > 0.05; 0 otherwise
Refinancing Size	= sale of common stock / net assets where refinancing = 1
Repurchases	= repurchases of common stock (prstkc less preferred repurchases (prstkpc))

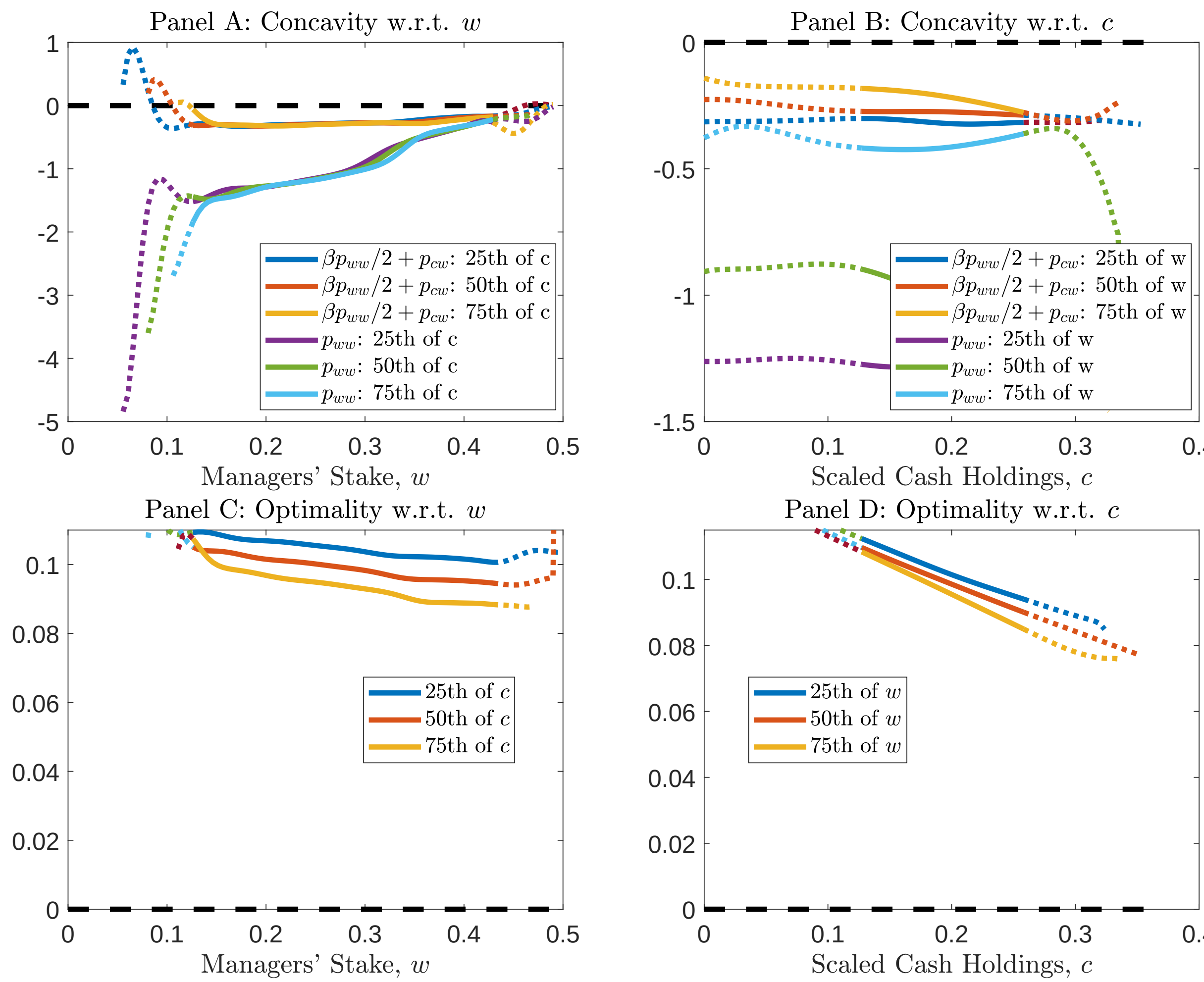


Figure A-1: Concavity and Optimality of the HJB Equation

Note: This figure displays in Panels A and B the concavity of the value function required for $\beta \geq \lambda(c)/(\mu(1-\tau_Y))$ to be the optimal solution and in Panels C and D the condition of full effort ($e_t = 1$) to be the optimal incentive strategy as we discuss in Appendix B. Panels A and B show p_{ww} and the sum $\beta p_{ww}/2 + p_{cw}$ with respect to w and c, respectively, across the 25th, 50th, and 75th percentiles of the marginal distribution of c and w, respectively. Panels C and D plot the value of the inequality that must to positive to ensure that full effort is preferred to a policy in which agents shirk ($e_t = 0$). The upper and lower 10 percent of the marginal distribution of w and c are dotted lines and the intermediate 10-90 percent are solid lines. The black dashed lines mark zero.

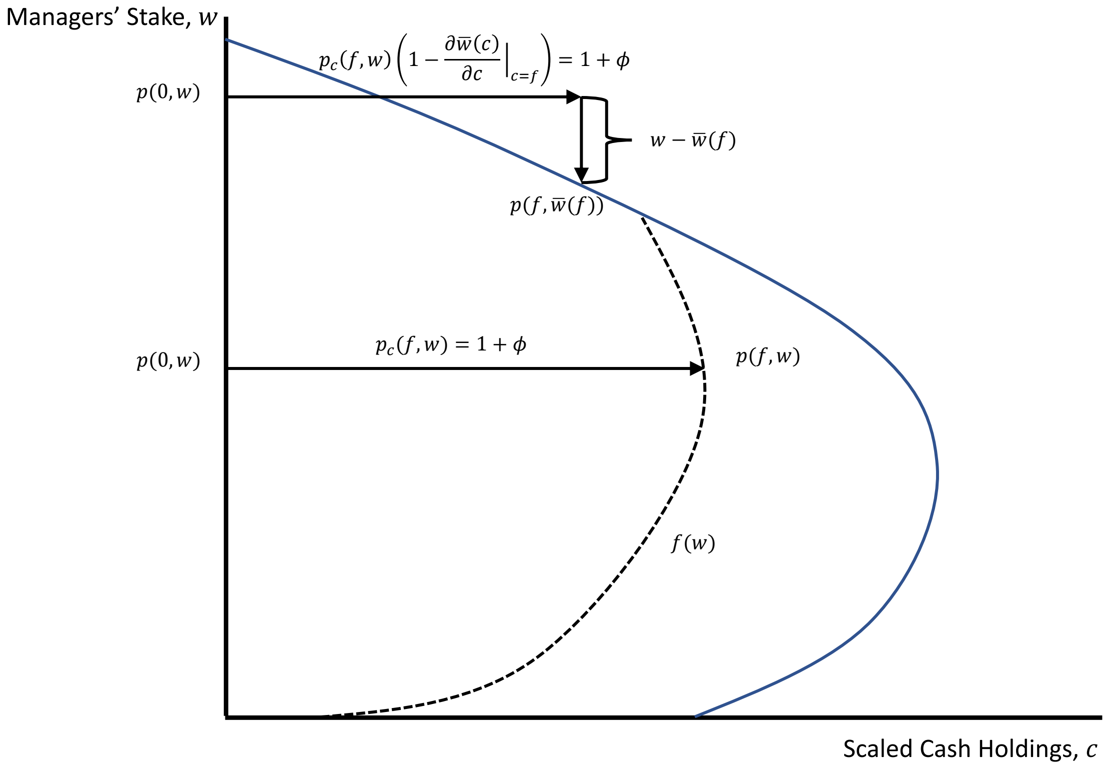


Figure A-2: Hypothetical Illustration of Alternative Setup

This figure hypothetically illustrates the refinancing decision and the state space of the alternative model where managers are paid out of the firm's cash holdings. The refinancing line is marked by the dashed line. The state space is the solid line.

Appendix C

Chapter 3 Appendices

A Strategies and general MPE definition

In this appendix we formally define (i) the class of default and debt adjustment strategies and (ii) the concepts of Subgame and Markov perfect equilibria.

Default and debt-adjustment strategies. Let I_t denote the debt adjustment process. As long as the firm faces a positive issuance cost $\beta > 0$, any debt adjustment process I_t whose paths have intervals of continuity leads to an infinite accumulation of adjustment costs. Building on this observation we define a default and adjustment strategy as a pair

$$\mathbf{s} \equiv \{\tau_b(\mathbf{s}), I(\mathbf{s})\}, \tag{A1}$$

where $\tau_b(\mathbf{s})$ is a stopping time that represents the time of default and $I_t(\mathbf{s})$ is a discrete process of the form

$$I_t(\mathbf{s}) = \sum_{s \in \mathcal{A}(\mathbf{s})} \mathbf{1}_{\{s \leq t\}} \Delta I_s(\mathbf{s}) = \sum_{s \in \mathcal{A}(\mathbf{s})} \mathbf{1}_{\{s \leq t\}} A_s(\mathbf{s}) F_{s-}, \tag{A2}$$

where $\mathcal{A}(\mathbf{s})$ is a thin set whose elements represent the moments at which the firm restructures its capital and $A_t(\mathbf{s}) \geq -1$ is a predictable process that represents the relative size of the adjustment conditional on a restructuring at date $t \geq 0$.

In most of the analysis, we will focus on Markov equilibria summarized by the variables F_t and Y_t that determine the cash flows of all stakeholders defined in equations (3.1) and (3.3). Accordingly, a strategy is said to be *Markovian* if

$$\begin{aligned} \mathcal{A}(\mathbf{s}) &= \{t \geq 0 : (F_{t-}, Y_t) \in \mathcal{R}\} \\ \tau_b(\mathbf{s}) &= \inf\{t \geq 0 : (F_t, Y_t) \in \mathcal{D}\}, \end{aligned} \tag{A3}$$

and

$$A_t(\mathbf{s}) = \mathbf{1}_{\{(F_{t-},Y_t)\in\mathcal{R}\}} A\left(F_{t-}, Y_t\right) \tag{A4}$$

for some closed disjoint subsets $\mathcal{D}, \mathcal{R}$ of $\mathbb{R}_+^2$ and some function $A = A(\cdot|\mathbf{s}) : \mathcal{R} \to [-1, \infty)$. If in addition

$$\begin{aligned} \mathcal{D} &= \left\{(F, Y) \in \mathbb{R}_+^2 : Y/F \in \bar{\mathcal{D}}\right\} \\ \mathcal{R} &= \left\{(F, Y) \in \mathbb{R}_+^2 : Y/F \in \bar{\mathcal{R}}\right\} \\ A\left(F, Y\right) &= a\left(Y/F\right) \end{aligned} \tag{A5}$$

for some closed disjoint subsets $\bar{\mathcal{D}}, \bar{\mathcal{R}}$ of $\mathbb{R}_+$ and some function $a = a(\cdot|\mathbf{s}) : \bar{\mathcal{R}} \to [-1, \infty)$, then we say that $\mathbf{s}$ is *reduced* Markovian. Throughout we denote by $\mathcal{S}_0$ the set of all default and adjustment strategies, by $\mathcal{M} \subset \mathcal{S}_0$ the subset of Markovian strategies, and by $\mathcal{M}_r \subset \mathcal{M}$ the subset of reduced Markovian strategies.

Markov Perfect Equilibrium. If creditors conjecture that the firm will use $\mathbf{a} \in \mathcal{S}_0$ but management instead uses another strategy $\mathbf{s} \in \mathcal{S}_0$, then the value of equity is

$$E_t(\mathbf{s}, \mathbf{a}) \equiv \mathbb{E}_t\left[\int_t^{\tau_b(\mathbf{s})} e^{-r(s-t)} \left(\delta(F_s, Y_s)\, ds + p_s(\mathbf{a})\, dI_s(\mathbf{s}) - \beta Y_s\, dN_s(\mathbf{s})\right)\right], \tag{A6}$$

subject to the cash flow dynamics (3.1) and

$$dF_s = -\xi F_{s-}\, ds + dI_s(\mathbf{s}), \tag{A7}$$

where

$$\delta(F_s, Y_s) \equiv (1-\tau)Y_s - \left(\xi + c(1-\tau)\right) F_s \tag{A8}$$

is the instantaneous cash flow to equity holders, $p_s(\mathbf{a})$ denotes the bond price, $dI_s(\mathbf{s}) = A_s(\mathbf{s})F_{s-}\, dN_s(\mathbf{s})$, and

$$N_t(\mathbf{s}) \equiv \sum_{s\in\mathcal{A}(\mathbf{s})} \mathbf{1}_{\{s\le t\}} \tag{A9}$$

is the counting process induced by the restructuring times of $\mathbf{s}$. To formalize the notion of an equilibrium, let $\mathcal{S}$ denote the set of feasible strategies such that

$$\Lambda(\mathbf{s}) \equiv \mathbb{E}\left[\int_0^{\tau_b(\mathbf{s})} e^{-rs} \left(F_s(\mathbf{s}) + Y_s\right)\, ds + \left(|\Delta F_s(\mathbf{s})| + Y_s\right)\, dN_s(\mathbf{s})\right] < \infty. \tag{A10}$$

The requirement that $\Lambda(\mathbf{s}) < \infty$ guarantees that all the expectations, and hence the value of all claims, are well defined.

Definition 4 (Subgame and Markov perfect equilibria). *A Subgame Perfect Equilibrium (SPE) is a strategy* $\mathbf{a} \in \mathcal{S}$ *such that*

$$E_t(\mathbf{a}, \mathbf{a}) = \sup_{\mathbf{s} \in \mathcal{S}} E_t(\mathbf{s}, \mathbf{a}), \quad t \geq 0. \tag{A11}$$

A Markov Perfect Equilibrium (MPE) is a SPE in the class $\mathcal{M} \subset \mathcal{S}$ *of Markov strategies. A Reduced Markov Perfect Equilibrium (rMPE) is a SPE in the class* $\mathcal{M}_r \subset \mathcal{S}$ *of reduced Markov strategies.*

The definition formalizes that in a subgame perfect equilibrium shareholders do not have incentive to deviate from the strategy $\mathbf{a}$ that creditors conjecture when pricing the bond. For ease of notation, in the main text we use the label MPE to denote both Markov and reduced-Markov perfect equilibria.

Lemma 14 in the Online Appendix, a version of the one-shot deviation principle, provides a characterization of SPEs in terms of a stochastic control problem in which the controlled process is two dimensional, (Y_t, F_t). Corollary 3 in the Online Appendix provides an alternative characterization that only involves a one-dimensional controlled process, $y_t = Y_t/F_t$.

B Propositions and proofs

Proposition 7 (Leverage ratchet effect). *If a feasible strategy* $\mathbf{a} \in \mathcal{S}$ *is a Markov Perfect Equilibrium, then the debt adjustment process* $I_t(\mathbf{a})$ *is non decreasing.*

Proof. Assume that $\mathbf{a} \in \mathcal{M}$ is a MPE in which $\mathbb{P}\left[\{s \in \mathcal{A}(\mathbf{a}) : dI_s(\mathbf{a}) < 0\}\right] \neq 0$. To show that this leads to a contradiction consider the deviation $\hat{\mathbf{a}} \in \mathcal{S}_0$ defined by the default time $\tau_b(\hat{\mathbf{a}}) \equiv \tau_b(\mathbf{a})$ and the face value process

$$F_t(\hat{\mathbf{a}}) \equiv \sup_{0 \leq u \leq t} \left\{ e^{\xi(u-t)} F_u(\mathbf{a}) \right\}. \tag{A12}$$

Standard results in the theory of Skorokhod reflection problems (see, e.g., Kruk, Lehoczky, Ramanan, and Shreve (2007) and references therein) show that we have

$$0 \leq \Delta I_t(\hat{\mathbf{a}}) = (F_t(\mathbf{a}) - F_{t-}(\hat{\mathbf{a}}))^+ \leq \Delta I_t(\mathbf{a})^+ \tag{A13a}$$

$$0 = (F_t(\mathbf{a}) - F_t(\hat{\mathbf{a}}))\, \Delta I_t(\hat{\mathbf{a}}). \tag{A13b}$$

Using these properties, we show in Lemma 15 in the Online Appendix that $\hat{\mathbf{a}} \in \mathcal{S}$ defines a feasible deviation and it remains to show that this deviation is profitable.

If bondholders conjecture that the firm will use a strategy $\mathbf{a} \in \mathcal{S}_0$ but shareholders instead use a different strategy $\mathbf{s} \in \mathcal{S}_0$, then the value of equity is

$$E_t(\mathbf{s}, \mathbf{a}) = \mathbb{E}_t \left[\int_t^{\tau_b(\mathbf{s})} e^{-r(s-t)} \left(\delta(F_s, Y_s)\, ds + p_s(\mathbf{a}) dI_s(\mathbf{s}) - \beta Y_s\, dN_s(\mathbf{s}) \right) \right], \tag{A14}$$

where $N_t(\mathbf{s}) \equiv \sum_{s \in \mathcal{A}(\mathbf{s})} \mathbf{1}_{\{s<t\}}$ is the counting process induced by the strategy $\mathbf{s}$ that keeps track of the debt restructuring time up to time t. Denote by $p(F,Y) = p(F,Y|\mathbf{a})$ the bond price function induced by the assumed Markov perfect equilibrium and by $\bar{c} \equiv c(1-\tau) + \xi$ the after tax cost of debt per unit of face value. A direct calculation using (A14) then shows

$$\begin{aligned}
\Delta E_0 \;&\equiv\; E_0(\hat{\mathbf{a}}, \mathbf{a}) - E_0(\mathbf{a}, \mathbf{a}) \\
&=\; \mathbb{E}\Bigg[\int_0^{\tau_b(\mathbf{a})} e^{-rs}\, \bar{c}\, (F_s(\mathbf{a}) - F_s(\hat{\mathbf{a}}))\, ds + \int_0^{\tau_b(\mathbf{a})} e^{-rs} \left(p(F_s(\hat{\mathbf{a}}), Y_s)\, dI_s(\hat{\mathbf{a}}) - \beta Y_s\, dN_s(\hat{\mathbf{a}}) \right) \\
&\qquad - \int_0^{\tau_b(\mathbf{a})} e^{-rs} \left(p(F_s(\mathbf{a}), Y_s)\, dI_s(\mathbf{a}) - \beta Y_s\, dN_s(\mathbf{a}) \right) \Bigg] \\
&\geq\; \mathbb{E}\left[\int_0^{\tau_b(\mathbf{a})} e^{-rs} \left(\bar{c} G_s\, ds - p\left(F_{s-}(\mathbf{a}), Y_s\right) \left(dG_s + \xi G_{s-}\, ds\right) \right) \right] \\
&=\; \mathbb{E}\left[\int_0^{\tau_b(\mathbf{a})} e^{-rs} (\bar{c} - \xi p\,(F_s(\mathbf{a}), Y_s)) G_s\, ds + \int_0^{\tau_b(\mathbf{a})} G_{s-}\, d\left(e^{-rs} p\,(F_s(\mathbf{a}), Y_s) \right) \right],
\end{aligned} \tag{A15}$$

where $G_t \equiv (F_t(\mathbf{a}) - F_t(\hat{\mathbf{a}}))$ is the difference between the face value processes associated with the two strategies, the inequality follows from (A13b) and the no jump condition $p\left(F_{s-}(\mathbf{a}), Y_s\right) = p\,(F_s(\mathbf{a}), Y_s)$, and the last equality follows from the fact that

$$p(F_0, Y_0|\mathbf{a})\, G_0 = G_{\tau_b(\mathbf{a})}\, p\Big(F_{\tau_b(\mathbf{a})}(\mathbf{a}), Y_{\tau_b(\mathbf{a})}\Big) = 0. \tag{A16}$$

Let $\rho \equiv r + \xi$. Since the process $e^{-\rho t} p(F_t(\mathbf{a}), Y_t) + \int_0^t e^{-\rho s}(c+\xi)\, ds$ is by construction a martingale on the stochastic interval $[\![0, \tau_b(\mathbf{a})]\!]$, we have that

$$d(p(F_t(\mathbf{a}), Y_t)) = (r+\xi) p(F_t(\mathbf{a}), Y_t)\, dt - (c+\xi) dt + dM_t \tag{A17}$$

for some local martingale M_t and therefore

$$d(e^{-rt} p\,(F_t(\mathbf{a}), Y_t)) = e^{-rt} \left(\xi p\,(F_{t-}(\mathbf{a}), Y_t) - (c+\xi) \right) dt + e^{-rt} dM_t. \tag{A18}$$

Substituting this dynamics into equation (A15) we obtain

$$E_0(\hat{\mathbf{a}}, \mathbf{a}) - E_0(\mathbf{a}, \mathbf{a}) \geq \mathbb{E}\left[\int_0^{\tau_b(\mathbf{a})} e^{-rs} G_{s-} dM_s - \int_0^{\tau_b(\mathbf{a})} e^{-rs} c \tau G_s\, ds \right], \tag{A19}$$

and the desired result now follows from Lemma 16 in the Online Appendix and the fact that G_t is non positive by construction. □

Proposition 8 (MPE characterization)**.** *A strategy $\mathbf{a} \in \mathcal{M}_r \cap \mathcal{S}$ is an rMPE if and only if the induced equity value $e(y|\mathbf{a})$ satisfies*

$$e(y|\mathbf{a}) \quad = \quad \sup_{\zeta \in \mathcal{T}} \mathbb{E}_y \left[\int_0^\zeta e^{-(r+\xi)t} \delta(\overline{y}_t)\, dt + e^{-(r+\xi)\zeta} \max\{\phi(\overline{y}_\zeta|\mathbf{a}), 0\} \right], \tag{A20}$$

where $\overline{y}_t$ is the uncontrolled scaled EBIT process

$$d\overline{y}_t = \overline{y}_t(\mu + \xi)dt + \overline{y}_t \sigma dW_t, \tag{A21}$$

$\delta(\overline{y}_t) = \delta(1, \overline{y}_t)$ is the instantaneous cash flow defined in equation (3.17)*, $\mathcal{T}$ denotes the set of stopping times, and $\phi(y|\mathbf{a})$ denotes the equity value upon restructuring,*

$$\phi(y|\mathbf{a}) \equiv \sup_{z \geq 0} \left\{ \frac{y}{z} e(z|\mathbf{a}) + \left(\frac{y}{z} - 1 \right) p(z|\mathbf{a}) - \beta y \right\}. \tag{A22}$$

In particular, if $\mathbf{a}$ is an MPE then (i) the induced scaled equity value is nonnegative, nondecreasing, convex, and differentiable at all points where $e(y|\mathbf{a}) = \max\{\phi(y|\mathbf{a}), 0\}$, and (ii) there exists a constant $0 \leq y_b(\mathbf{a}) < \infty$ such that $e(y|\mathbf{a}) = 0 = \max\{\phi(y|\mathbf{a}), 0\}$ at all points $y \leq y_b(\mathbf{a})$.

Proof. Equation (A20) follows from Lemma 17 by noting that

$$\begin{aligned} \frac{R(\bar{F}_t, Y_t|\mathbf{a})}{\bar{F}_t} &= \sup_{z \geq 0} \left\{ \frac{Y_t}{z\bar{F}_t} e\,(z|\mathbf{a}) + \left(\frac{Y_t}{z\bar{F}_t} - 1 \right) p\,(z|\mathbf{a}) \right\} - \frac{\beta Y_t}{z\bar{F}_t} \\ &= \sup_{z \geq 0} \left\{ \frac{\bar{x}_t}{z} e\,(z|\mathbf{a}) + \left(\frac{\bar{x}_t}{z} - 1 \right) p\,(z|\mathbf{a}) \right\} - \beta \bar{x}_t \\ &= \phi(\bar{x}_t|\mathbf{a}), \end{aligned} \tag{A23}$$

and $\bar{F}_t = e^{-\xi t} F_0$. Equation (A20) can equivalently be written as follows

$$e(y|\mathbf{a}) \quad = \quad \hat{e}(y) + \sup_{\zeta \in \mathcal{T}} \mathbb{E}_y \left[e^{-\rho\zeta} \psi(\overline{y}_\zeta|\mathbf{a}) \right], \tag{A24}$$

with

$$\hat{e}(y) = \mathbb{E}_y \left[\int_0^\infty e^{-\rho t} \delta(\overline{y}_t)\, dt \right] = \frac{\delta(0)}{\rho} + \frac{\delta(y) - \delta(0)}{r - \mu}, \tag{A25}$$

and

$$\psi(y|\mathbf{a}) \equiv \phi(y|\mathbf{a})^+ - \hat{e}(y). \tag{A26}$$

To see that equation (A20) is equivalent to (A24) it suffices to observe that the no-action equity value function satisfies the Dynkin identity

$$\hat{e}(y) - \mathbb{E}_y \left[e^{-\rho\zeta} \hat{e}(\bar{y}_\zeta) \right] = \mathbb{E}_y \left[\int_0^\zeta e^{-\rho t} \delta(\bar{y}_t) dt \right] \tag{A27}$$

for all stopping times $\zeta \in \mathcal{T}$. Setting $\theta \equiv 0$ in (A20) shows that the equity value function is nonnegative. On the other hand, we have that $\psi(y|\mathbf{a})$ is convex as the supremum of a family of affine functions and it thus follows from Alvarez et al. (2003, Theorem 5) and Lamberton, Zervos, et al. (2013, Corollary 7.5) that $v(y) \equiv e(y|\mathbf{a}) - \hat{e}(y)$ is differentiable at all points of the set

$$\{y \geq 0 : v(y|\mathbf{a}) = \psi(y|\mathbf{a})\} = \{y \geq 0 : e(y|\mathbf{a}) = \phi(y|\mathbf{a})^+\}. \tag{A28}$$

Since the function $\hat{e}(y)$ is linear this in turn implies that $e(y|\mathbf{a})$ is also convex and differentiable at all points of this set.

Finally note that since $e(y|\mathbf{a}) \geq \hat{e}(y)$ we have $e(y|\mathbf{a}) > 0$ for all sufficiently large y and thus $\mathcal{D}_r(\mathbf{a}) \neq \mathbb{R}_+$. Let $y_b(\mathbf{a}) \equiv \sup\{y \geq 0 : y \in \mathcal{D}_r(\mathbf{a})\}$. Since the scaled equity value function is nonnegative and not identically zero, then $y_b(\mathbf{a}) < \infty$ and $e'_+(z|\mathbf{a}) > 0$ at some point $z > y_b(\mathbf{a})$. Together with the convexity proved above, this implies that the scaled equity value is nondecreasing and it follows by continuity that $e(y|\mathbf{a}) = 0 \geq \phi(y|\mathbf{a})^+$ for all $y \leq y_b(\mathbf{a})$. □

Corollary 2 (Restructuring). *If $\mathbf{a} \in \mathcal{M}_r \cap \mathcal{S}$ is a rMPE then for all $y \in \bar{\mathcal{R}}(\mathbf{a})$ the equity value is given by*

$$e(y|\mathbf{a}) \quad = \quad \sup_{z \in \mathcal{I}_r(\mathbf{a})} \Phi(y, z|\mathbf{a}), \tag{A29}$$

where $\Phi(y, z|\mathbf{a})$ is defined in equation (A22) *and $\mathcal{I}_r(\mathbf{a}) \equiv \mathbb{R}_+ \setminus \bar{\mathcal{D}}(\mathbf{a}) \cup \bar{\mathcal{R}}(\mathbf{a})$ denotes the inaction region. The set of target income-to-debt ratio upon restructuring, $\left\{\mathcal{Y}(y) \equiv \frac{y}{(1+a(y|\mathbf{a}))}\right\}$, is given by*

$$\{\mathcal{Y}(y)\} = \operatorname*{argmax}_{z \in \mathcal{I}_r(\mathbf{a})} \Phi(y, z|\mathbf{a}). \tag{A30}$$

Furthermore, the scaled equity value is differentiable and satisfies the envelope condition

$$e'(y|\mathbf{a}) = \frac{\partial \Phi}{\partial y} e(y, \mathcal{Y}(y)|\mathbf{a}) \tag{A31}$$

for all points in the restructuring region, $y \in \bar{\mathcal{R}}(\mathbf{a})$.

Proof. If $y \in \bar{\mathcal{R}}(\mathbf{a})$ lies then it follows from equation (A20) that

$$e(y|\mathbf{a}) \geq \phi(y|\mathbf{a})^+ = \sup_{z \geq 0} \Phi(y, z|\mathbf{a})^+ \tag{A32}$$

and from equation (A11) in the Online Appendix that

$$e(y|\mathbf{a}) = \Phi(y, \mathcal{Y}(y)|\mathbf{a}) = \frac{y}{\mathcal{Y}(y)} e(\mathcal{Y}(y)|\mathbf{a}) + \left(\frac{x}{\mathcal{Y}(y)} - 1\right) p(\mathcal{Y}(y)|\mathbf{a}) - \beta y. \tag{A33}$$

Combining the two equations shows that we have

$$\begin{aligned}
&\bar{\mathcal{R}}(\mathbf{a}) \subseteq \{y \geq 0 : e(y|\mathbf{a}) = \phi(y|\mathbf{a}) \geq 0\} \\
&\mathcal{Y}(y) \in \mathcal{Z} = \operatorname*{argmax}_{z\geq 0} \Phi\left(y, z|\mathbf{a}\right),
\end{aligned} \tag{A34}$$

and the first part will follow if we can show that the maximizer is unique and lies in $\mathcal{I}_r(\mathbf{a})$. Suppose to the contrary that $y \in \bar{\mathcal{R}}(\mathbf{a})$ is such that

$$\sup_{z\geq 0} \Phi(y, z|\mathbf{a}) = \Phi(y, z^*|\mathbf{a}). \tag{A35}$$

for some $z^* \notin \mathcal{I}_r(\mathbf{a})$. If $z^* \in \bar{\mathcal{D}}(\mathbf{a})$ then it follows from (A34) that we have

$$e(y|\mathbf{a}) = \Phi(y, z^*|\mathbf{a}) = -\beta y < 0, \tag{A36}$$

which contradicts the nonnegativity of the scaled equity value function. On the other hand, if $z^* \in \bar{\mathcal{R}}(\mathbf{a})$ then

$$\begin{aligned}
e(y|\mathbf{a}) &= \Phi(y, z^*|\mathbf{a}) \\
&= \frac{y}{z^*} e(z^*|\mathbf{a}) + \left(\frac{y}{z^*} - 1\right) p(z^*|\mathbf{a}) - \beta y \\
&= \frac{y}{z^*} \Phi(z^*, \mathcal{Y}(z^*)|\mathbf{a}) + \left(\frac{y}{z^*} - 1\right) p(z^*|\mathbf{a}) - \beta y \\
&= \Phi(y, \mathcal{Y}(z^*)) + \left(\frac{y}{z^*} - 1\right) \left(p(z^*|\mathbf{a}) - p(\mathcal{Y}(z^*)|\mathbf{a})\right) - \beta y \\
&= \Phi(y, \mathcal{Y}(z^*)) - \beta y < \Phi(y, \mathcal{Y}(z^*)),
\end{aligned} \tag{A37}$$

where the third equality follows from equation (A11) in the Online Appendix, the fifth equality follows from the no jump condition $p(z^*|\mathbf{a}) = p(\mathcal{Y}(z^*)|\mathbf{a})$, and the inequality follows from the strict positivity of the fixed cost. This contradicts the fact that $e(y|\mathbf{a}) = \phi(y|\mathbf{a})$ over $\bar{\mathcal{R}}(\mathbf{a})$ and thus establishes that $\mathcal{Z} \subseteq \mathcal{I}_r(\mathbf{a})$. To complete the proof, observe that

$$e(y|\mathbf{a}) = \phi(y|\mathbf{a}) = \sup_{z\in\mathcal{I}_r(\mathbf{a})} \Phi(y, z|\mathbf{a}) \tag{A38}$$

is differentiable at all points of $\bar{\mathcal{R}}(\mathbf{a})$ as a result of equation (A34) and Proposition 8, and apply Milgrom and Segal (2002, Corollary 4.iii)). $\square$

The following proposition provides necessary condition for an MPE in barrier strategies:

Proposition 9 (Necessary conditions for barrier-strategy MPE)**.** *Assume that the barrier strategy $\mathbf{a} = (y_b, y_u, \mathcal{Y}(y))$ is an MPE. Then the following conditions are satisfied:*

(i) Default boundary: $y_b < y_0$, with $y_0 \equiv \frac{\Pi}{\Pi-1}\frac{r-\mu}{\rho}\left(c + \frac{\xi}{1-\tau}\right)$ and Π given in equation (3.15)*.*

(ii) Limited liability: $e(y|\mathbf{a}) = \max\{\phi(y|\mathbf{a}), 0\} = 0$ *for* $y \in (0, y_b]$, *where* $\phi(y|\mathbf{a})$ *is the equity continuation value defined in equation* (A22).

(iii) Equity valuation in the restructuring region: $e(y|\mathbf{a}) = \max\{\phi(y|\mathbf{a}), 0\} > 0$ *for* $y \in [y_u, \infty)$.

(iv) Value-matching and smooth-pasting at the default boundary y_b*:*

$$e(y_b|\mathbf{a}) = e'(y_b|\mathbf{a}) = 0. \tag{A39}$$

(v) Value-matching and smooth-pasting at the restructuring boundary y_u*:*

$$e(y|\mathbf{a}) = \frac{y}{\mathcal{Y}(y)} e(\mathcal{Y}(y)|\mathbf{a}) + \left(\frac{y}{\mathcal{Y}(y)} - 1\right) p(\mathcal{Y}(y)|\mathbf{a}) - \beta y, \quad y \geq y_u \tag{A40}$$

$$e'(y|\mathbf{a}) = \frac{e(\mathcal{Y}(y)|\mathbf{a}) + p(\mathcal{Y}(y)|\mathbf{a})}{\mathcal{Y}(y)} - \beta, \quad y \geq y_u. \tag{A41}$$

(vi) Optimality of restructuring:

$$\{\mathcal{Y}(y)\} = \underset{z \in (y_b, y_u)}{\operatorname{argmax}} \left\{ \frac{y}{z} e(z|\mathbf{a}) + \left(\frac{y}{z} - 1\right) p(z|\mathbf{a}) - \beta y \right\}, \quad y \geq y_u. \tag{A42}$$

Proof of (i) If $\mathbf{a}$ is an MPE with $y_b(\mathbf{a}) > y_0$, then $e(y|\mathbf{a}) = 0 < e_0(y)$ for all $y \in (y_0, y_b(\mathbf{a}))$ which contradicts equation (A11) of the Online Appendix.

Proof of (ii) If $\mathbf{a} \in \mathcal{B}$ is an MPE then it follows from Proposition 8 and the definition of the strategy that we have $e(y|\mathbf{a}) = 0 \geq \phi(y|\mathbf{a})^+$ for all $y \leq y_b(\mathbf{a})$.

Proof of (iv) Since by definition $e(y|\mathbf{a}) = 0$ for all $y \leq y_b(\mathbf{a})$ it follows from Proposition 8 that $e(y|\mathbf{a}) = \phi(y|\mathbf{a})^+ = 0$ over that region. Hence, the scaled equity value function is differentiable at all points $x \leq y_b(\mathbf{a})$ and the desired result follows by noting that $e'_-(y|\mathbf{a}) = 0$ at any such point.

Proof of (iii) Since by definition $e(y|\mathbf{a}) = \Phi(y, \mathcal{Y}(y)|\mathbf{a}) \leq \phi(y|\mathbf{a})$ for $y \geq y_u(\mathbf{a})$ it follows from Proposition 8 that we have

$$0 \leq e(y|\mathbf{a}) = \Phi(y, \mathcal{Y}(y)|\mathbf{a}) = \phi(y|\mathbf{a}), \qquad y \geq y_u(\mathbf{a}). \tag{A43}$$

To see that the inequality is strict note that due to Item (iv) the scaled equity value function solves (3.24) subject to (3.25)–(3.26). In particular,

$$\lim_{y \downarrow y_b(\mathbf{a})} \frac{1}{2}\sigma^2 x^2 e''(y|\mathbf{a}) = -\delta(y_b(\mathbf{a})) > 0, \tag{A44}$$

where the strict inequality follows from Item (i) and the definition of the no-issuance default threshold. The above inequality implies that we have $e(y|\mathbf{a}) > 0$ in a right neighborhood of $y_b(\mathbf{a})$ and thus for all $y > y_b(\mathbf{a})$ by convexity.

Proof of (vi) This follows directly from Corollary 2.

Proof of (v) Since $e(y|\mathbf{a}) = \phi(y|\mathbf{a}) > 0$ for all $y \geq y_u(\mathbf{a})$ by Item (iii) it follows from Proposition 8 that the scaled equity value function, and thus also $\phi(y|\mathbf{a})$, is differentiable at all points $y \geq y_u(\mathbf{a})$. On the other hand, by Item (vi) we have that $\mathcal{Y}(y)$ is the unique maximizer of the function $z \mapsto \Phi(y, z|\mathbf{a})$ over the compact set $\mathcal{I}_r(a)$ and the validity of (A41) now follows from Milgrom and Segal (2002, Corollary 4) and (3.26). □

Proposition 10 (Properties of equity in an MPE)**.** *Assume that* $\mathbf{a}$ *is a barrier strategy such that* $\delta(y_b(\mathbf{a})) < 0$ *and that satisfies conditions (i) and (iv) of Proposition 9. Then*

(i) $e(y|\mathbf{a})$ *is nonnegative, nondecreasing, convex on the interval* $[0, y_u)$ *and strictly positive on the interval* (y_b, y_u)*.*

(ii) $e(y|\mathbf{a}) \geq e_0(y)$ *for all* $y \leq y_u$ *if and only if* $y_b \leq y_0$*, where* $y_0 \equiv y_{b,Leland}$ *and* $e_0(y) \equiv e_{Leland}(y)$ *defined in equations* (3.45) *and* (3.47)*, respectively.*

Proof of (i) Since $e(y|\mathbf{a})$ solves (3.24) subject to value matching and smooth pasting at the default boundary the uniqueness of the solution to second order differential equations implies that the constants in equation (3.27) can be expressed as

$$M_e(\mathbf{a}) = \frac{y_b(\mathbf{a})^{-\Theta}(1-\tau)\,(y_b(\mathbf{a}) - y_0)\,(\Pi - 1)}{(r-\mu)(\Theta - \Pi)} \geq 0$$

$$N_e(\mathbf{a}) = \frac{y_b(\mathbf{a})^{-\Pi}(1-\tau)(y_b(\mathbf{a})(\Theta-1)\Pi + y_0\Theta(1-\Pi))}{(r-\mu)\Pi(\Pi-\Theta)} \geq 0. \tag{A45}$$

Therefore, $e(y|\mathbf{a})$ is convex on the interval $[0, y_u(\mathbf{a}))$ and the remaining claims in the statement follow by observing that, because

$$\lim_{y \downarrow y_b(\mathbf{a})} \frac{1}{2}\sigma^2 y^2 e''(y) = -\delta(y_b(\mathbf{a})) > 0, \tag{A46}$$

we must have $\min\{e, e'\}(y|\mathbf{a}) > 0$ in a right neighborhood of $y_b(\mathbf{a})$ and thus over the whole interval since the scaled equity value is convex.

Proof of (ii) The necessity of the condition is clear since in its absence $e(y|\mathbf{a}) = 0 < e_0(y)$ for all $y \leq (y_0, y_b(\mathbf{a}))$. Now assume that $y_b(\mathbf{a}) \leq y_0$. If $y_u(\mathbf{a}) \leq y_0$ then the result follows from Item i) since $e_0(y) = 0$ on $[0, y_0]$. Assume from now on that $y_u(\mathbf{a}) > y_0$. Proceeding as in the first part of the proof shows that

$$w(x) = e(y|\mathbf{a}) - e_0(y) = M_e(\mathbf{a})x^{\Theta} + \mathbf{1}_{\{x > y_0\}}\bar{N}(\mathbf{a})x^{\Pi}, \qquad x \in [y_0, y_u(\mathbf{a})), \tag{A47}$$

where $M_e(\mathbf{a}) \geq 0$ and

$$\bar{N}(\mathbf{a}) \equiv \frac{y_0^{1-\Pi}(1-\tau)}{(r-\mu)\Pi} + \frac{y_b(\mathbf{a})^{-\Pi}(1-\tau)(y_b(\mathbf{a})(\Theta-1)\Pi + y_0\Theta(1-\Pi))}{(r-\mu)(\Pi-\Theta)\Pi}. \tag{A48}$$

Noting that $\bar{N}(\mathbf{a})\big|_{y_{b(\mathbf{a})=y_0}} = 0$ and

$$\frac{d\bar{N}(\mathbf{a})}{dy_b(\mathbf{a})} = \frac{(1-\tau)(\Pi-1)(y_b(\mathbf{a})(\Theta-1) - y_0\Theta)}{y_b(\mathbf{a})^{1+\Pi}(r-\mu)(\Theta-\Pi)} \geq 0, \qquad y_b(\mathbf{a}) \leq y_0, \tag{A49}$$

we deduce that $\bar{N}(\mathbf{a}) \leq 0$. This implies that $w(x)$ is non decreasing on $[y_0, y_u(\mathbf{a})]$ and the thesis follows by observing that $w(y_0) = e(y_0|\mathbf{a}) \geq 0$. $\square$

Proof of Proposition 4

By construction we have that

$$\begin{aligned} p(y|\mathbf{a}) &= p(\mathcal{Y}(y)|\mathbf{a}), \qquad y \geq y_u \\ e(y|\mathbf{a}) &= \Phi(y, \mathcal{Y}(y)|\mathbf{a}) = \sup_{z\in[x_b, y_u]} \Phi(y, z|\mathbf{a}), \qquad y \geq y_u \end{aligned} \tag{A50}$$

and, since the scaled equity value function is differentiable at y_u, we deduce from Proposition 10.i) that the function $e(y|\mathbf{a})$ is globally convex, non decreasing, and strictly positive on the interval $[y_b, \infty)$. A direct calculation using the above expressions shows that

$$\Phi(y, z|\mathbf{a}) = \Phi(y, \mathcal{Y}(z)|\mathbf{a}) - \beta y < \Phi(y, \mathcal{Y}(y)|\mathbf{a}), \qquad z \geq y_u \tag{A51}$$

and the first claim follows by observing that $\Phi(y, z|\mathbf{a}) = -y\beta$ for all $z \leq y_b$. Now, since the function $e(y|\mathbf{a})$ is convex it follows from Proposition 12 in the Online Appendix that $\mathbf{a}$ is an equilibrium if and only if the convex function

$$v(y) \equiv e(y|\mathbf{a}) - \hat{e}(y) \tag{A52}$$

is a weak solution to the HJB equation (A76). On the interval $[0, y_b]$ we have that $v(y) = -\hat{e}(y)$ and

$$\mathcal{O}v(dy) = \delta(y_b)dy \leq \delta(y_0)dy < 0, \tag{A53}$$

so that $v(y)$ is a weak solution on that interval if and only if equation (3.35) holds for all $x \leq x_b$. On the interval (x_b, y_u) we have that $\mathcal{O}v(dy) \equiv 0$ and it follows that $v(y)$ is a weak solution on that interval if and only if equation (3.35) holds for $x \in [x_b, y_u]$. Finally, since $e(y|\mathbf{a}) > 0$ for $y > y_b$ we have that

$$v(y) = e(y|\mathbf{a}) - \hat{e}(y) = \Phi(y, \mathcal{Y}(y)|\mathbf{a}) - \hat{e}(y) = \psi(y|\mathbf{a}) \tag{A54}$$

for all $y \in [y_u, \infty)$ and it follows that $v(y)$ is a weak solution on that interval if and only if the restriction of the measure

$$\mathcal{O}v(dy) = \mathcal{O}e(dy|\mathbf{a}) + \mathcal{O}\hat{e}(dy) = \mathcal{O}e(dy|\mathbf{a}) + \delta(y)dy \tag{A55}$$

to that interval is non positive. To complete the proof we now provide an expression for this measure. First observe that as a result of Condition (A42) and Milgrom and Segal (2002, Corollary 4) we have that

$$e(y|\mathbf{a}) = \Phi(y, \mathcal{Y}(y)|\mathbf{a}) = \sup_{z \in [x_b, y_u]} \Phi(y, z|\mathbf{a}) \tag{A56}$$

is continuously differentiable at all points of $[y_u, \infty)$ and satisfies

$$e'(y|\mathbf{a}) = s(\mathcal{Y}(y)|\mathbf{a}) - \beta = s(y|\mathbf{a}) = \frac{1}{y}\left(e(y|\mathbf{a}) + p(\mathcal{Y}(y)|\mathbf{a})\right), \qquad y \geq y_u, \tag{A57}$$

where the last two equalities equality follow from (A50) and

$$s(y|\mathbf{a}) \equiv \frac{1}{y}\left(e(y|\mathbf{a}) + p(y|\mathbf{a})\right), \tag{A58}$$

denotes the enterprise value of the firm per unit of cash flow. Since $e(y|\mathbf{a})$ is convex, $\mathcal{Y}(y)$ is continuous on $x \geq y_u$, and $p(y|\mathbf{a})$ is continuous on $[0, y_u]$ we have this derivative is continuous as well as non decreasing and, therefore, absolutely continuous. Combining this property with equation (A57) we obtain

$$e''(dy|\mathbf{a}) = s'(y|\mathbf{a})dy = \frac{1}{y}\left(e'(y|\mathbf{a}) + p'(y|\mathbf{a}) - s(y|\mathbf{a})\right)dy = \frac{1}{y}p'(y|\mathbf{a})dy. \tag{A59}$$

Using equations (A57) and (A59), it follows that the non positivity of equation (A55) is equivalent to the supermartingale condition in equation (3.39):

$$\begin{aligned} 0 \;&\geq\; \delta(y) - (r-\mu)\, e(y|\mathbf{a}) + (\mu+\xi)\, p(y|\mathbf{a}) + \frac{1}{2}\sigma^2 y\, p'(y|\mathbf{a}) \\ &=\; \delta(y) - (r+\xi)\, e(y|\mathbf{a}) + (\mu+\xi) y\, e'(y|\mathbf{a}) + \frac{\sigma^2}{2} y^2\, e''(y|\mathbf{a}). \end{aligned} \tag{A60}$$

When $F = 0$, the value of equity given restructuring amount $\hat{F}$ is:

$$E(0, Y) = E(\hat{F}, Y) + P(\hat{F}, Y) - \beta Y = \hat{F}\left(e(\hat{y}) + p(\hat{y}) - \beta\hat{y}\right), \tag{A61}$$

where, for ease of notation, we have suppressed the dependence of the debt and equity value on the strategy $\mathbf{a}$. Therefore:

$$\frac{E(0, Y)}{Y} = \frac{\hat{F}}{Y}\left(e(\hat{y}) + p(\hat{y}) - \beta\hat{y}\right) = \frac{e(\hat{y}) + p(\hat{y})}{\hat{y}} - \beta. \tag{A62}$$

This implies that choosing the best restructuring amount $\hat{F}$ (or, equivalently $\hat{y}$) determines the maximum scaled equity value at the restructuring time, that is,

$$\frac{E(0,Y)}{Y} = \lim_{F\to 0}\frac{E(F,Y)}{Y} = \lim_{F\to 0}\frac{E(F,Y)/F}{Y/F} = \lim_{y\to\infty}\frac{e(y)}{y} = \sup_{\hat{y}}\left\{\frac{e(\hat{y})}{\hat{y}} + \frac{p(\hat{y})}{\hat{y}} - \beta\right\}, \quad \text{(A63)}$$

The above condition implies that the equity value in the limit for $y \to \infty$ is

$$e(y) = \sup_{z}\Phi_\infty(y,z) = \sup_{z\geq 0}\left\{\frac{y}{z}e(z) + \frac{y}{z}p(z) - \beta y\right\} = \Phi_\infty(y,\mathcal{Y}(y)). \quad \text{(A64)}$$

From equation (A64), at the point $y = \infty$, the equity value is

$$e(y) = \Phi_\infty(y,\mathcal{Y}(y)) = \frac{y}{\mathcal{Y}(y)}e(\mathcal{Y}(y)) + \frac{y}{\mathcal{Y}(y)}p(\mathcal{Y}(y)) - \beta y, \quad \text{(A65)}$$

where $\mathcal{Y}(y) = \mathrm{argmax}_{z\in[y_b,y_u]}\,\Phi_\infty(y,z)$. For the measure (A55) to be non positive we need

$$\mathcal{L}e(y) - (r+\xi)e(y) + \delta(y) \leq 0, \quad \text{for} \quad y \to \infty. \quad \text{(A66)}$$

Taking the derivative of $e(y) = \Phi_\infty(y,\mathcal{Y}(y))$ with respect of y we have

$$e'(y) = \frac{\partial\Phi_\infty(y,\mathcal{Y}(y))}{\partial y} + \underbrace{\frac{\partial\Phi_\infty(y,\mathcal{Y}(y))}{\partial\mathcal{Y}(y)}}_{=0 \text{ by Envelope Theorem}}\frac{\partial\mathcal{Y}(y)}{\partial y} = \frac{\partial\Phi_\infty(y,\mathcal{Y}(y))}{\partial y}. \quad \text{(A67)}$$

Using the expression of $\Phi_\infty(x,\mathcal{Y}(y))$ in equation (A65) we have

$$\Phi_\infty(y,\mathcal{Y}(y)) = \frac{y}{\mathcal{Y}(y)}e(\mathcal{Y}(y)) + \frac{y}{\mathcal{Y}(y)}p(\mathcal{Y}(y)) - \beta y. \quad \text{(A68)}$$

By direct calculation we obtain that at $y \to \infty$

$$e'(y) = \frac{e(y)}{y} \quad \text{and} \quad e''(y) = 0. \quad \text{(A69)}$$

Substituting equations (A69) in equation (A66) we obtain

$$\delta(y) - (r-\mu)e(y) \leq 0, \quad y \to \infty. \quad \text{(A70)}$$

□

Proof of Proposition 5

In the inaction region $y_t \in (y_b, y_u)$, debt satisfies the ODE (3.10) subject to the boundary conditions (3.54)–(3.55), while equity satisfies the ODE (3.24), subject to the boundary conditions (3.56)–(3.57). Imposing these boundary conditions allows us determine the constants M_p, N_p, M_e, and N_e as a function of the barrier strategy parameters $\mathbf{a}_{\mathrm{GJL}} = (y_b, y_u, \hat{y})$. Since shareholders cannot commit to a default threshold y_b before issuing debt, y_b is determined by the smooth pasting condition $e'(y_b) = 0$. In contrast, commitment to future restructuring implies that the restructuring threshold y_u and target $\hat{y}$ is determined by the first-order conditions (3.51). □

Proof of Corollary 1

To identify the limiting case $\beta \to 0$, we consider a Taylor series expansion around the point $\epsilon \equiv y_u/\hat{y} - 1 > 0$. We can then express $p(\hat{y})$ as follows

$$p(\hat{y}) = p\left(\frac{y_u}{1+\epsilon}\right) \approx p(y_u) - y_u p'(y_u)\epsilon + o(\epsilon^2). \tag{A71}$$

Therefore, as $\beta \to 0$ the upper boundary condition for debt in equation (3.55) can be written as $p'(y_u) = 0$, which is condition (3.58). Similarly, setting $\beta \to 0$ in the equity boundary condition (3.59) and considering a Taylor series expansion around the point $\epsilon \equiv y_u/\hat{y} - 1 > 0$ we obtain

$$\begin{aligned} e(y_u) &= \frac{y_u}{\hat{y}} e(\hat{y}) + \left(\frac{y_u}{\hat{y}} - 1\right) p(\hat{y}) = (1+\epsilon) e\left(\frac{y_u}{1+\epsilon}\right) + \epsilon p\left(\frac{y_u}{1+\epsilon}\right) \\ &\approx (1+\epsilon)\left(e(y_u) - y_u e'(y_u)\epsilon\right) + \epsilon\left(p(y_u) - y_u p'(y_u)\epsilon\right) + o(\epsilon^2). \end{aligned} \tag{A72}$$

Simplifying and ignoring terms of order $o(\epsilon^2)$ we obtain that as $\beta \to 0$ condition (3.57) converges to condition (3.59). $\square$

Proof of Proposition 6

It is convenient to express asset values in terms of the prices of two other securities: one that pays $e^{-\xi(\tau_b - t)}$ the first time y_b is reached, conditional upon $\tau_b < \tau_u$, and one that pays $e^{-\xi(\tau_u - t)}$ the first time y_u is reached, conditional upon $\tau_b > \tau_u$, that is,

$$\begin{aligned} \pi_b(y_t) &= \mathbb{E}_t\left[e^{-(r+\xi)(\tau_b - t)}\mathbf{1}_{(\tau_b < \tau_u)}\right] = \frac{\left(\frac{y_t}{y_u}\right)^{\Pi} - \left(\frac{y_t}{y_u}\right)^{\Theta}}{\left(\frac{y_b}{y_u}\right)^{\Pi} - \left(\frac{y_b}{y_u}\right)^{\Theta}} \\ \pi_u(y_t) &= \mathbb{E}_t\left[e^{-(r+\xi)(\tau_u - t)}\mathbf{1}_{(\tau_u < \tau_b)}\right] = \frac{\left(\frac{y_t}{y_b}\right)^{\Theta} - \left(\frac{y_t}{y_b}\right)^{\Pi}}{\left(\frac{y_u}{y_b}\right)^{\Theta} - \left(\frac{y_u}{y_b}\right)^{\Pi}}, \end{aligned} \tag{A73}$$

with $\Theta > 1$ and $\Pi < 0$ constants defined in equations (3.14) and (3.15). Let $e_0(y|y_b)$ be the equity value when the firm chooses not to issue new debt in the future (i.e., $y_u = \infty$) and to default at a boundary y_b, that is,

$$e_0(y_t|y_b) = \hat{e}(y_t) - \frac{1-\tau}{r-\mu}\left(\frac{y_t}{y_0}\right)^{\Pi}\left[y_0\left(\frac{1-\Pi}{\Pi}\right) + y_b\right], \tag{A74}$$

where $\hat{e}(y)$ is defined in equation (3.28) and y_0 is given in equation (3.45). We can then write the claim to equity as:

$$
\begin{aligned}
e(y_t|\mathbf{a}) &= e_0(y_t|y_b) + \pi_u(y_t)\left[e(y_u|\mathbf{a}) - e_0(y_u|y_b)\right] \\
&= e_0(y_t|y_b) + \left(\left(\frac{y_t}{y_b}\right)^{\Theta} - \left(\frac{y_t}{y_b}\right)^{\Pi}\right)\left[\frac{e(y_u|\mathbf{a}) - e_0(y_u|y_b)}{\left(\frac{y_u}{y_b}\right)^{\Theta} - \left(\frac{y_u}{y_b}\right)^{\Pi}}\right]. \qquad \text{(A75)}
\end{aligned}
$$

where $\mathbf{a} = (y_b, y_u, \hat{y})$. The term $e_0(y_t|y_b) - \pi_u(y_t)\, e_0(y_u|y_b)$ reflects the present value of the claim to dividends for all dates $t < \tau$ prior y reaching either boundary, and the term $\pi_u(y_t)e(y_u|\mathbf{a})$ captures the present value for which equity could be sold for at $y = y_u$.

Note that in equation (A75) the policy parameters $(y_u, \hat{y})$ appear only in the term in square brackets and that this term is independent of y_t. Therefore, for a given y_b, the optimal choices for $(y_u^*(y_b), \hat{y}^*(y_b))$ are independent of the value of y_t:

$$
\begin{aligned}
\frac{\partial e(y_t|\mathbf{a})}{\partial y_u} &= \frac{\partial}{\partial y_u}\left[\frac{e(y_u|\mathbf{a}) - e_0(y_u|y_b)}{\left(\frac{y_u}{y_b}\right)^{\Theta} - \left(\frac{y_u}{y_b}\right)^{\Pi}}\right] = 0 \\
\frac{\partial e(y_t;\theta)}{\partial \hat{y}} &= \frac{\partial e(y_u|\mathbf{a})}{\partial \hat{y}} = 0. \qquad \text{(A76)}
\end{aligned}
$$

This implies that, for any given y_b, and for *all* values of $y_t \in (y_b, y_u)$, equity values are maximized by choosing the optimal policy parameters $(y_u^*(y_b), \hat{y}^*(y_b))$ determined from equations (A76)-(A76). We note that $e'(y_b)$ is an increasing function of y_b. Thus, for values of y_b below the optimal value y_b^*, $e'(y_b) < 0$, implying that the limited liability condition is not satisfied. In contrast, for values of y_b above the optimal value, $e'(y_b) > 0$.

Note that if $e'(y_b) > 0$, and hence, $y_b > y_b^*$, then the claim to equity is increased for *all* values of $y_t \in (y_b, y_u)$ by lowering the default boundary by a small amount $\epsilon y_b > 0$, that is,

$$
e'(y_b) > 0 \quad \Rightarrow \quad e(y_t|y_b(1-\epsilon)) \; > \; e(y_t|y_b) \;\; \text{for all} \;\; y_t, \qquad \text{(A77)}
$$

which, expressed in differential form, implies that $\frac{\partial e(y_t|\mathbf{a})}{\partial y_b} < 0$. Furthermore, Lemma 19 in the Online Appendix shows that when y_b is chosen to satisfy the smooth pasting condition, the equity claim is increasing and convex in y.

Therefore it follows that the optimal policy with commitment, which is identified by imposing the smooth pasting condition $e'(y_b) = 0$, and the first order conditions in equation (3.51), is in fact the global optimal barrier policy subject to limited liability. $\square$

Appendix D

Chapter 3 Online Appendices

This Online Appendix contains additional results needed for the formal characterization of a Markov Perfect Equilibrium (MPE).

A Existence of MPE: Auxiliary results

Lemma 14 (Characterization of SPE)**.** *A strategy* $\mathbf{a} \in \mathcal{S}$ *is a* SPE *if and only if*

$$E_t(\mathbf{a}, \mathbf{a}) = \sup_{\mathbf{s} \in \mathcal{S}} \mathbb{E}_t \Bigg[\int_t^{\theta_t(\mathbf{s}) \wedge \tau_b(\mathbf{s})} e^{-r(s-t)} \, \delta(F_s, Y_s) \, ds \tag{A1}$$

$$+ \mathbf{1}_{\{\theta_t(\mathbf{s}) < \tau_b(\mathbf{s})\}} \, e^{-r(\theta_t(\mathbf{s})-t)} \left(E_{\theta_t(\mathbf{s})}(\mathbf{a}, \mathbf{a}) + p_{\theta_t(\mathbf{s})}(\mathbf{a}) \Delta I_{\theta_t(\mathbf{s})}(\mathbf{s}) - \beta Y_{\theta_t(\mathbf{s})} \right) \Bigg], \tag{A2}$$

where the stopping time

$$\theta_t(\mathbf{s}) \equiv \inf\{s \geq t : s \in \mathcal{A}(\mathbf{s})\} = \inf\{s \geq t : dI_s(\mathbf{s}) \neq 0\} \tag{A3}$$

denotes the time of the first restructuring prescribed by the strategy $\mathbf{s} \in \mathcal{S}$ *on or after an arbitrary date* $t \geq 0$*.*

Proof. Assume that $\mathbf{a} \in \mathcal{S}$ is a SPE, let $\mathbf{s} \in \mathcal{S}$ and consider for each fixed starting point $t \geq 0$ the one-shot deviation $\mathbf{s}_t$ obtained by following $\mathbf{s}$ until $\tau_b(\mathbf{s}) \wedge \theta_t(\mathbf{s})$ and then reverting to $\mathbf{a}$. Using the equilibrium property of $\mathbf{a}$ together with the law of iterated expectations we deduce that

$$E_t\left(\mathbf{a},\mathbf{a}\right)=\mathbb{E}_t\Bigg[\int_t^{\theta_t(\mathbf{a})\wedge\tau_b(\mathbf{a})} e^{-r(s-t)}\delta(F_s,Y_s)\,ds \tag{A4}$$

$$+\mathbf{1}_{\{\theta_t(\mathbf{a})<\tau_b(\mathbf{a})\}}e^{-r(\theta_t(\mathbf{a})-t)}\left(E_{\theta_t(\mathbf{a})}(\mathbf{a},\mathbf{a})+p_{\theta_t(\mathbf{a})}(\mathbf{a})\Delta I_{\theta_t(\mathbf{a})}(\mathbf{a})-\beta Y_{\theta_t(\mathbf{a})}\right)\Bigg] \tag{A5}$$

$$\geq E_t\left(\mathbf{s}_t,\mathbf{a}\right)=\mathbb{E}_t\Bigg[\int_t^{\theta_t(\mathbf{s})\wedge\tau_b(\mathbf{s})} e^{-r(s-t)}\delta(F_s,Y_s)\,ds \tag{A6}$$

$$+\mathbf{1}_{\{\theta_t(\mathbf{s})<\tau_b(\mathbf{s})\}}e^{-r(\theta_t(\mathbf{s})-t)}\left(E_{\theta_t(\mathbf{s})}(\mathbf{a},\mathbf{a})+p_{\theta_t(\mathbf{s})}(\mathbf{a})\Delta I_{\theta_t(\mathbf{s})}(\mathbf{s})-\beta Y_{\theta_t(\mathbf{s})}\right)\Bigg] \tag{A7}$$

and the necessity of (A1) follows from the arbitrariness of $\mathbf{s}\in\mathcal{S}$. To establish the converse assume that $\mathbf{a}\in\mathcal{S}$ satisfies (A1). Since never restructuring and defaulting at the first time that the cash flow becomes negative is feasible we have that $E_t(\mathbf{a},\mathbf{a})\geq 0$ at all times. Using this property and iterating (A1) forward we deduce that

$$E_t(\mathbf{a},\mathbf{a})\geq\mathbb{E}_t\left[\int_t^{\tau_b(\mathbf{s})\wedge\theta_{n,t}(\mathbf{s})} e^{-r(s-t)}\left(\delta(F_s,Y_s)\,ds+p_s(\mathbf{a})dI_s(\mathbf{s})-\beta Y_s\,dN_s(\mathbf{s})\right)\right] \tag{A8}$$

where $\theta_{n,t}(\mathbf{s})$ is the time of the nth restructuring after $t\geq 0$. let $Z_{n,t}$ denote the random variable inside the conditional expectation. Since the bond price is bounded and $\delta(F,Y)$ is a linear function we have that

$$\sup_{n\geq 1}|Z_{n,t}|\leq C_0(t)\int_0^{\tau_b(\mathbf{s})} e^{-rs}\left((F_s+Y_s)\;ds+(|\Delta I_s(\mathbf{s})|+Y_s)\,dN_s(\mathbf{s})\right) \tag{A9}$$

for some deterministic function $C_0(t)>0$ and it follows from (A10) that the right hand side has finite expectation. Letting the number of restructuring rounds $n\to\infty$ and appealing to the dominated convergence theorem then gives

$$E_t(\mathbf{a},\mathbf{a})\geq\mathbb{E}_t\left[\int_t^{\tau_b(\mathbf{s})} e^{-r(s-t)}\left(\delta(F_s,Y_s)\,ds+p_s(\mathbf{a})dI_s(\mathbf{s})-\beta Y_s\,dN_s(\mathbf{s})\right)\right] \tag{A10}$$

and the desired result now follows from (A14). □

Corollary 3. *A default and adjustment strategy* $\mathbf{a}\in\mathcal{S}$ *is a* SPE *if and only if the scaled equity value process* $e_t(\mathbf{a})=E_t(\mathbf{a},\mathbf{a})/F_t$ *satisfies*

$$e_t(\mathbf{a})=\sup_{\mathbf{s}\in\mathcal{S}}\mathbb{E}_t\Bigg[\int_t^{\theta_t(\mathbf{s})\wedge\tau_b(\mathbf{s})} e^{-\rho(s-t)}\delta(y_s)\,ds$$

$$+\mathbf{1}_{\{\theta_t(\mathbf{s})<\tau_b(\mathbf{s})\}}e^{-\rho(\theta_t(\mathbf{s})-t)}\Big((1+A_{\theta_t(\mathbf{s})}(\mathbf{s}))e_{\theta_t(\mathbf{s})}(\mathbf{a})+p_{\theta_t(\mathbf{s})}(\mathbf{a})A_{\theta_t(\mathbf{s})}(\mathbf{s})-\beta y_{\theta_t(\mathbf{s})-}\Big)\Bigg] \tag{A11}$$

with the discount rate $\rho=r+\xi$ *and the cash flow function* $\delta(y)\equiv\delta(1,y)$.

Proof. The result follows from Lemma 14 by noting that we have

$$F_s = e^{-\xi(s-t)} F_t, \qquad \text{for all } s \in [t, \theta_t(\mathbf{s})), \tag{A12}$$

and therefore

$$\begin{aligned} & E_{\theta_t(\mathbf{s})}(\mathbf{a},\mathbf{a}) + p_{\theta_t(\mathbf{s})}(\mathbf{a})\Delta I_{\theta_t(\mathbf{s})}(\mathbf{s}) - \beta Y_{\theta_t(\mathbf{s})} & \text{(A13)}\\ &\quad = F_{\theta_t(\mathbf{s})-}\left(\frac{F_{\theta_t(\mathbf{s})}}{F_{\theta_t(\mathbf{s})-}} e_{\theta_t(\mathbf{s})}(\mathbf{a}) + p_{\theta_t(\mathbf{s})}(\mathbf{a})\frac{\Delta I_{\theta_t(\mathbf{s})}(\mathbf{s})}{F_{\theta_t(\mathbf{s})-}} - \beta y_{\theta_t(\mathbf{s})-}\right) & \text{(A14)}\\ &\quad = F_{\theta_t(\mathbf{s})-}\Big(\big(1 + A_{\theta_t(\mathbf{s})}(\mathbf{s})\big)\, e_{\theta_t(\mathbf{s})}(\mathbf{a}) + p_{\theta_t(\mathbf{s})}(\mathbf{a}) A_{\theta_t(\mathbf{s})}(\mathbf{s}) - \beta y_{\theta_t(\mathbf{s})-}\Big) & \text{(A15)}\\ &\quad = e^{-\xi(\theta_t(\mathbf{s})-t)} F_t\Big(\big(1 + A_{\theta_t(\mathbf{s})}(\mathbf{s})\big)\, e_{\theta_t(\mathbf{s})}(\mathbf{a}) + p_{\theta_t(\mathbf{s})}(\mathbf{a}) A_{\theta_t(\mathbf{s})}(\mathbf{s}) - \beta y_{\theta_t(\mathbf{s})-}\Big) & \text{(A16)} \end{aligned}$$

where the second equality follows from the definition of $A_t(\mathbf{s}) \geq -1$ as the relative size of the debt adjustment. ☐

Lemma 15. *The strategy* $\hat{\mathbf{a}}$ *considered in Proposition 7 is feasible, that is,* $\hat{\mathbf{a}} \in \mathcal{S}$.

Proof. Using equations (A13) we deduce that the deviation $\hat{\mathbf{a}}$ satisfies $dN_t(\hat{\mathbf{a}}) \leq dN_t(\mathbf{a})$ as well as $|\Delta F_t(\hat{\mathbf{a}})| \leq |\Delta F_t(\mathbf{a})|$. To see this, note that the face value $\hat{\mathbf{a}}$ jumps only if the face value $\mathbf{a}$ jumps to a new maximum. Therefore, either $dN(\mathbf{a}) = 1$ and $dN(\hat{\mathbf{a}}) = 0$ or $dN(\mathbf{a}) = dN(\hat{\mathbf{a}})$, and by the same token either $|\Delta F_t(\mathbf{a})| > 0$ and $\Delta F_t(\hat{\mathbf{a}}) = 0$ or $|\Delta F_t(\mathbf{a})| > 0$ and $\Delta F_t(\hat{\mathbf{a}}) > 0$ but in this case

$$\begin{aligned} \Delta F_t(\hat{\mathbf{a}}) &= F_t(\hat{\mathbf{a}}) - F_{t-}(\hat{\mathbf{a}}) \leq F_t(\hat{\mathbf{a}}) - F_{t-}(\mathbf{a}) & \text{(A17)}\\ &= F_t(\mathbf{a}) - F_{t-}(\mathbf{a}) = \Delta F_t(\mathbf{a}) & \text{(A18)} \end{aligned}$$

because $F_t(\mathbf{a}) \leq F_t(\hat{\mathbf{a}})$ by construction and the two face values processes are reset to a common level upon a jump of the $\hat{\mathbf{a}}$ face value. From $dN_t(\hat{\mathbf{a}}) \leq dN_t(\mathbf{a})$ and $|\Delta F_t(\hat{\mathbf{a}})| \leq |\Delta F_t(\mathbf{a})|$ it follows that

$$\mathbb{E}\left[\int_0^{\tau_b(\mathbf{a})} e^{-rs}\left(\left(|\Delta F_s(\hat{\mathbf{a}})| + Y_s\right) dN_s(\hat{\mathbf{a}}) - \left(|\Delta F_s(\mathbf{a})| + Y_s\right) dN_s(\mathbf{a})\right)\right] \leq 0. \tag{A19}$$

On the other hand, Itô's formula implies that we have

$$F_s(\hat{\mathbf{a}}) - F_s(\mathbf{a}) = \int_0^s e^{\xi(u-s)}\left(dI_u(\hat{\mathbf{a}}) - dI_u(\mathbf{a})\right) \tag{A20}$$

and therefore

$$\int_0^{\tau_b(\mathbf{a})} e^{-rs}\left(F_s(\hat{\mathbf{a}}) - F_s(\mathbf{a})\right) ds = \int_0^{\tau_b(\mathbf{a})} ds\, e^{-\rho s}\left\{\int_0^s e^{\xi u}\left(dI_u(\hat{\mathbf{a}}) - dI_u(\mathbf{a})\right)\right\} \tag{A21}$$

$$\leq \int_0^{\tau_b(\mathbf{a})} ds\, e^{-\rho s}\left\{\int_0^s e^{\xi u}|\Delta F_u(\mathbf{a})|\, dN_u(\mathbf{a})\right\} \tag{A22}$$

$$= \int_0^{\tau_b(\mathbf{a})} e^{-ru}\, dN_u(\mathbf{a})|\Delta F_u(\mathbf{a})|\left\{\int_u^{\tau_b(\mathbf{a})} e^{-\rho(s-u)}\, ds\right\} \tag{A23}$$

$$\leq \frac{1}{\rho}\int_0^{\tau_b(\mathbf{a})} e^{-ru}|\Delta F_u(\mathbf{a})|\, dN_u(\mathbf{a}). \tag{A24}$$

where the first inequality follows from (A13b), that is,

$$\begin{aligned} dI_u(\hat{\mathbf{a}}) - dI_u(\mathbf{a}) &= -\mathbf{1}_{\{dN_u(\hat{\mathbf{a}})=0\}} dI_u(\mathbf{a}) \\ &= -\mathbf{1}_{\{dN_u(\hat{\mathbf{a}})=0\}} \Delta F_u(\mathbf{a}) dN_u(\mathbf{a}) \\ &\leq \mathbf{1}_{\{dN_u(\hat{\mathbf{a}})=0\}} |\Delta F_u(\mathbf{a})| dN_u(\mathbf{a}) \\ &\leq |\Delta F_u(\mathbf{a})| dN_u(\mathbf{a}). \end{aligned} \tag{A25}$$

Combining (A19) and (A24) then shows that $\Lambda(\hat{\mathbf{a}}) \leq C_0\Lambda(\mathbf{a})$ for some $C_0 > 0$ and the desired result follows. □

Lemma 16. *The process*

$$U_t \equiv \int_0^{t\wedge\tau_b(\mathbf{a})} e^{-rs} G_{s-} dM_s \tag{A26}$$

that appears in equation (A19) *of Proposition 7 has expected value zero.*

Proof. Denote by $p_t(\mathbf{a}) = p(F_t(\mathbf{a}), Y_t)$ the bond price along the path of the assumed equilibrium. From

$$d\left(e^{-rt}p_t(\mathbf{a})\right) = e^{-rt}\xi p_t(\mathbf{a})dt - e^{-rt}(c+\xi)dt + e^{-rt}dM_t \tag{A27}$$

and

$$dG_t = -\xi G_t dt + dI_t(\mathbf{a}) - dI_t(\hat{\mathbf{a}}) \tag{A28}$$

we obtain

$$\begin{aligned} d(e^{-rt}G_t p_t(\mathbf{a})) &= G_{t-}p(e^{-rt}p_t(\mathbf{a})) + e^{-rt}p_{t-}(\mathbf{a})dG_t \\ &= e^{-rt}\xi G_t p(\mathbf{a})dt - e^{-rt}G_t(c+\xi)dt + e^{-rt}G_{t-}dM_t \\ &+ e^{-rt}p_{t-}(\mathbf{a})[-\xi G_t dt + dI_t(\mathbf{a}) - dI_t(\hat{\mathbf{a}})] \\ &= e^{-rt}G_{t-}dM_t - e^{-rt}G_t(c+\xi)dt + e^{-rt}p_{t-}(\mathbf{a})[dI_t(\mathbf{a}) - dI_t(\hat{\mathbf{a}})] \end{aligned} \tag{A29}$$

where the first equality follows from the fact that G and p have neither common jumps nor common exposure to the BM. Integrating over $[0,\theta]$, with $\theta \equiv \tau_b(\mathbf{a}) \wedge t$ and recalling that $G_0 = 0$ we get:

$$\begin{aligned}\int_0^\theta e^{-rs} G_{s-} dM_{s-} e^{-r\theta} G_\theta p_\theta(\mathbf{a}) = & \int_0^\theta e^{-rs}\left(G_s(c+\xi)\,ds - p_{s-}(\mathbf{a})(dI_t(\mathbf{a}) - dI_t(\hat{\mathbf{a}}))\right) \\ = & \int_0^\theta e^{-rs}\left(G_s(c+\xi)\,ds - p_s(\mathbf{a})(dI_t(\mathbf{a}) - dI_t(\hat{\mathbf{a}}))\right) \quad \text{(A30)}\end{aligned}$$

where the last equality follows from the no jump condition. Using the definition of U_t in equation (A26), we can then write

$$U_t = e^{-r\theta} G_\theta p_\theta(\mathbf{a}) + \int_0^\theta e^{-rs}\left((c+\xi)G_s\,ds - p_s(\mathbf{a})\left(dI_s(\mathbf{a}) - dI_s(\hat{\mathbf{a}})\right)\right)\Bigg|_{\theta\equiv\tau_b(\mathbf{a})\wedge t} \quad \text{(A31)}$$

and it thus follows from the uniform boundedness of the bond price process, the non positivity of G_t and (A13a) that we have that, for some constant C_1

$$\begin{aligned}\frac{|U_t|}{C_1} &\leq e^{-rt\wedge\tau_b(\mathbf{a})}|G_{t\wedge\tau_b(\mathbf{a})}| + \int_0^{t\wedge\tau_b(\mathbf{a})} e^{-rs}\left((c+\xi)|G_s|\,ds + |dI_s(\mathbf{a}) - dI_s(\hat{\mathbf{a}})|\right) \\ &= \int_0^{t\wedge\tau_b(\mathbf{a})} e^{-rs}\left((r-c)G_s\,ds + dI_s(\hat{\mathbf{a}}) - dI_s(\mathbf{a}) + |dI_s(\mathbf{a}) - dI_s(\hat{\mathbf{a}})|\right) \\ &\leq \int_0^{\tau_b(\mathbf{a})} e^{-rs}\left((F_s(\hat{\mathbf{a}}) + F_s(\mathbf{a}))\,|r-c|\,ds + |\Delta F_s(\mathbf{a})| dN_s(\mathbf{a})\right) \quad \text{(A32)}\end{aligned}$$

where the first equality follows by noting that, since $F_t(\mathbf{a}) \leq F_t(\hat{\mathbf{a}})$, $G_t \leq 0$ and therefore

$$\begin{aligned}p(e^{-rt}|G_t|) &= -p(e^{-rt}G_t) & \text{(A33)} \\ &= e^{-rt}((r+\xi)G_t dt - dI_t(\mathbf{a}) + dI_t(\hat{\mathbf{a}})) & \text{(A34)}\end{aligned}$$

and

$$e^{-rT}|G_T| = \int_0^T e^{-rt}((r+\xi)G_t dt - dI_t(\mathbf{a}) + dI_t(\hat{\mathbf{a}})) \quad \text{(A35)}$$

Substituting into the first inequality in equation (A32) gives the stated equality. To prove the second inequality in (A32) note that

$$(r-c)G_t \leq |r-c||F_t(\mathbf{a}) - F_t(\hat{\mathbf{a}})| \leq |r-c|(F_t(\mathbf{a}) + F_t(\hat{\mathbf{a}})). \quad \text{(A36)}$$

Using equation (A25)

$$\begin{aligned}&|dI_t(\mathbf{a}) - dI_t(\hat{\mathbf{a}})| - (dI_t(\mathbf{a}) - dI_t(\hat{\mathbf{a}})) & \text{(A37)} \\ &\quad = \mathbf{1}_{\{dN_t(\hat{\mathbf{a}})=0\}}|dI_t(\mathbf{a})| \leq |dI_t(\mathbf{a})| & \text{(A38)}\end{aligned}$$

the second inequality in (A32). This in turn implies that

$$\mathbb{E}\left\{\sup_{t\geq 0}|U_t|\right\} \leq C_2\left(\Lambda(\mathbf{a})+\Lambda(\hat{\mathbf{a}})\right) < \infty \tag{A39}$$

for some constant $C_2 > 0$ where the second inequality follows from the fact that $\mathbf{a}$ and $\hat{\mathbf{a}}$ are both feasible by Lemma 15. This shows that the local martingale U_t is a uniformly integrable martingale and the desired result follows. □

The following lemma shows that the search for Markov equilibria is equivalent to solving a recursive optimal stopping problem.

Lemma 17. *A Markovian strategy* $\mathbf{a} \in \mathcal{M} \cap \mathcal{S}$ *is a* MPE *if and only if the induced equity value function satisfies*

$$E(F, Y|\mathbf{a}) = \sup_{\theta\in\mathcal{T}} \mathbb{E}_{F,Y}\left[\int_0^\theta e^{-rt}\delta(\bar{F}_t, Y_t)dt + e^{-r\theta}R\left(\bar{F}_\theta, Y_\theta|\mathbf{a}\right)^+\right] \tag{A40}$$

subject to (3.1) *and the uncontrolled dynamics*

$$d\bar{F}_t = -\xi\bar{F}_t dt \tag{A41}$$

where the reward function is defined by

$$R(F, Y|\mathbf{a}) \equiv \sup_{G\in\mathbb{R}_+}\left\{E(G, Y|\mathbf{a}) + (G-F)\,p(G, Y|\mathbf{a}) - \beta Y\right\} \tag{A42}$$

and $\mathcal{T}$ *denotes the set of stopping times.*

Proof of necessity. Assume that $\mathbf{a} \in \mathcal{M} \cap \mathcal{S}$ is a MPE and denote by

$$R(F, Y, G|\mathbf{a}) \equiv E(G, Y|\mathbf{a}) + (G-F)\,p(G, Y|\mathbf{a}) - \beta Y \tag{A43}$$

the objective function on the right hand side of (A42). Since $(\tau_b(\mathbf{a}), \theta_0(\mathbf{a}))$ are stopping times it follows from (A42) and Lemma 14 that

$$\begin{aligned}
E(F, Y|\mathbf{a}) = \mathbb{E}_{F,Y}\Bigg[&\int_0^{\tau_b(\mathbf{a})\wedge\theta_0(\mathbf{a})} e^{-rt}\delta(\bar{F}_t, Y_t)dt &\text{(A44)}\\
&+ e^{-r\theta_0(\mathbf{a})}\mathbf{1}_{\{\theta_0(\mathbf{a})<\tau_b(\mathbf{a})\}}R\left(\bar{F}_{\theta_0(\mathbf{a})}, Y_{\theta_0(\mathbf{a})}, \bar{F}_{\theta_0(\mathbf{a})}\left(1+A(\bar{F}_{\theta_0(\mathbf{a})}, Y_{\theta_0(\mathbf{a})})\right)\Big|\,\mathbf{a}\right)\Bigg] &\text{(A45)}\\
&\leq \sup_{(\tau,\theta)\in\mathcal{T}^2}\mathbb{E}_{F,Y}\left[\int_0^{\tau\wedge\theta} e^{-rt}\delta(\bar{F}_t, Y_t)dt + \mathbf{1}_{\{\theta<\tau\}}e^{-r\theta}R\left(\bar{F}_\theta, Y_\theta|\mathbf{a}\right)\right] &\text{(A46)}\\
&\leq \sup_{(\tau,\theta)\in\mathcal{T}^2}\mathbb{E}_{F,Y}\left[\int_0^{\tau\wedge\theta} e^{-rt}\delta(\bar{F}_t, Y_t)dt + \mathbf{1}_{\{\theta<\tau\}}e^{-r\theta}R\left(\bar{F}_\theta, Y_\theta|\mathbf{a}\right)^+\right] &\text{(A47)}\\
&\leq \sup_{\zeta\in\mathcal{T}}\mathbb{E}_{F,Y}\left[\int_0^{\zeta} e^{-rt}\delta(\bar{F}_t, Y_t)dt + e^{-r\zeta}R\left(\bar{F}_\zeta, Y_\zeta|\mathbf{a}\right)^+\right] &\text{(A48)}
\end{aligned}$$

To establish the reverse inequality let

$$R_n(F, Y|\mathbf{a}) \equiv \sup_{0 \le G \le n} R(F, Y, G|\mathbf{a}) \tag{A49}$$

and consider the sequence $(\mathbf{s}_n)_{n=1}^{\infty}$ of one shot deviations defined by

$$\theta_0(\mathbf{s}_n) \equiv \sigma + \mathbf{1}_{\{R_n(\bar{F}_\sigma, Y_\sigma|\mathbf{a}) \le 0\}} \infty, \tag{A50}$$

$$\tau_{_b}(\mathbf{s}_n) \equiv \mathbf{1}_{\{R_n(\bar{F}_\sigma, Y_\sigma|\mathbf{a}) \le 0\}} \sigma + \mathbf{1}_{\{R_n(\bar{F}_\sigma, Y_\sigma|\mathbf{a}) > 0\}} (\sigma + q_\sigma \circ \tau_{_b}(\mathbf{a})), \tag{A51}$$

and

$$\bar{F}_{\theta_0(\mathbf{s}_n)} \left(1 + A_{\theta_0(\mathbf{s}_n)}(\mathbf{s}_n)\right) = \operatorname*{argmax}_{0 \le G \le n} R\left(\bar{F}_{\theta_0(\mathbf{s}_n)}, Y_{\theta_0(\mathbf{s}_n)}, G|\mathbf{a}\right) \tag{A52}$$

where σ is an arbitrary but fixed stopping time, and q_σ denotes the Markov shift operator. It is easily seen that $\mathbf{s}_n \in \mathcal{S}$ is a feasible deviation for each $n \ge 1$. Therefore, it follows from Lemma 14 and the specification of $\mathbf{s}_n$ that we have

$$E(F, Y|\mathbf{a}) \ge \mathbb{E}_{F,Y} \left[\int_0^{\tau_{_b}(\mathbf{s}_n) \wedge \theta_0(\mathbf{s}_n)} e^{-rt} \delta(\bar{F}_t, Y_{_t}) dt \right. \tag{A53}$$

$$\left. + e^{-r\theta_0(\mathbf{s}_n)} \mathbf{1}_{\{\theta_0(\mathbf{s}_n) < \tau_{_b}(\mathbf{s}_n)\}} R_n(\bar{F}_{\theta_0(\mathbf{s}_n)}, Y_{\theta_0(\mathbf{s}_n)}|\mathbf{a}) \right] \tag{A54}$$

$$= \mathbb{E}_{F,Y} \left[\int_0^{\sigma} e^{-rt} \delta(\bar{F}_t, Y_{_t}) dt + e^{-r\sigma} R_n(\bar{F}_\sigma, Y_\sigma|\mathbf{a})^+ \right]. \tag{A55}$$

Letting $n \to \infty$ on both sides and invoking the monotone convergence theorem to justify the interchange of limit and expectation then gives

$$E(F, Y|\mathbf{a}) \ge \mathbb{E}_{F,Y} \left[\int_0^{\sigma} e^{-rt} \delta(\bar{F}_t, Y_{_t}) dt + e^{-r\sigma} R(\bar{F}_\sigma, Y_\sigma|\mathbf{a})^+ \right] \tag{A56}$$

and the result follows by taking the supremum over $\sigma \in \mathcal{T}$.

Proof of sufficiency. Assume that $\mathbf{a} \in \mathcal{M} \cap \mathcal{S}$ satisfies (A40) and let $\mathbf{s} \in \mathcal{S}$ be fixed. Because $\tau_b(\mathbf{s}) \wedge \theta_t(\mathbf{s})$ is a stopping time this implies that we have

$$E(F_t, Y_t|\mathbf{a}) \geq \mathbb{E}_t\Bigg[\int_t^{\tau_b(\mathbf{s})\wedge\theta_t(\mathbf{s})} e^{-r(s-t)}\delta(\bar{F}_s, Y_s)\, ds \tag{A57}$$

$$+ e^{-r\left(\tau_b(\mathbf{s})\wedge\theta_t(\mathbf{s})-t\right)} R\left(\bar{F}_{\tau_b(\mathbf{s})\wedge\theta_t(\mathbf{s})}, Y_{\tau_b(\mathbf{s})\wedge\theta_t(\mathbf{s})}\Big|\,\mathbf{a}\right)^+\Bigg] \tag{A58}$$

$$\geq \mathbb{E}_t\Bigg[\int_t^{\tau_b(\mathbf{s})\wedge\theta_t(\mathbf{s})} e^{-r(s-t)}\delta(\bar{F}_s, Y_s)\, ds \tag{A59}$$

$$+ e^{-r(\theta_t(\mathbf{s})-t)}\mathbf{1}_{\{\theta_t(\mathbf{s})<\tau_b(\mathbf{s})\}} R\left(\bar{F}_{\theta_t(\mathbf{s})}, Y_{\theta_t(\mathbf{s})}\big|\,\mathbf{a}\right)\Bigg] \tag{A60}$$

$$\geq \mathbb{E}_t\Bigg[\int_t^{\tau_b(\mathbf{s})\wedge\theta_t(\mathbf{s})} e^{-r(s-t)}\delta(\bar{F}_s, Y_s)\, ds \tag{A61}$$

$$+ e^{-r(\theta_t(\mathbf{s})-t)}\mathbf{1}_{\{\theta_t(\mathbf{s})<\tau_b(\mathbf{s})\}} R\left(\bar{F}_{\theta_t(\mathbf{s})}, Y_{\theta_t(\mathbf{s})}, \bar{F}_{\theta_t(\mathbf{s})}\left(1 + A_{\theta_t(\mathbf{s})}(\mathbf{s})\right)\big|\,\mathbf{a}\right)\Bigg] \tag{A62}$$

and the required result now follows from Lemma 14, the arbitrariness of $\mathbf{s} \in \mathcal{S}$ and the definition of the function $R(F, Y, G|\mathbf{a})$. □

Lemma 18. *Barrier strategies are feasible, that is, $\mathcal{B} \subseteq \mathcal{S}$.*

Proof. Fix a barrier strategy $\mathbf{a} \in \mathcal{B}$. Since the pair (F_t, Y_t) forms a Markov process we have that $\Lambda(\mathbf{a}) = \Lambda(F_0, Y_0)$ for some (possibly infinite) function $\Lambda : \mathbb{R}_+^2 \to \mathbb{R} \cup \{\infty\}$ that satisfies the boundary conditions

$$\Lambda(F, Y) = 0, \qquad (F, Y) \in \mathcal{D}(\mathbf{a}), \tag{A63}$$

$$\Lambda(F, Y) = Y\left(1 + \frac{1}{\mathcal{Y}(y)}\right) + \Lambda\left(\frac{Y}{\mathcal{Y}(y)}, Y\right), \qquad (F, Y) \in \mathcal{R}(\mathbf{a}). \tag{A64}$$

On the other hand, a standard calculation using Girsanov's theorem and the law of iterated expectations shows that

$$\Lambda(F, Y) = \lambda(y)Y, \qquad (F, Y) \in \mathbb{R}_+\backslash(\mathcal{D} \cup \mathcal{R})(\mathbf{a}) \tag{A65}$$

for some function $\lambda : \mathbb{R}_+ \to \mathbb{R} \cup \{\infty\}$ that satisfies

$$\lambda(y) = G(y) + H(y)\left(1 + \frac{1}{\mathcal{Y}(y_u(\mathbf{a}))} - \frac{1}{x} + \lambda(\mathcal{Y}(y_u(\mathbf{a})))\right) \tag{A66}$$

with $H, G : \mathbb{R}_+ \to \mathbb{R}_+$ uniformly bounded and such that

$$\min\{G(y), 1 - H(y)\} > 0, \quad y \in \mathcal{I}_r(\mathbf{a}) \equiv (y_b(\mathbf{a}), y_u(\mathbf{a})). \tag{A67}$$

Combining (A63), (A64), and (A66) we deduce that the strategy is feasible if and only if the constant

$$\lambda(\mathcal{Y}(y_u(\mathbf{a}))) = \frac{G(\mathcal{Y}(y_u(\mathbf{a}))) + H(\mathcal{Y}(y_u(\mathbf{a})))}{1 - H(\mathcal{Y}(y_u(\mathbf{a})))} \tag{A68}$$

is finite and the desired result now follows from (A67) since the point $\mathcal{Y}(y_u(\mathbf{a}))$ lies by assumption in the set $\mathcal{I}_r(\mathbf{a})$. □

Lemma 19. *The equity value $e_{SP}(y|\mathbf{a})$ corresponding to a barrier strategy $\mathbf{a} = (y_b, y_u, \hat{y})$ in which y_b is chosen to satisfy the smooth pasting condition $e'(y_b) = 0$ is increasing and convex in y.*

Proof. Let us define $\Gamma(\mathbf{a})$ as

$$\Gamma(\mathbf{a}) = e(y_u|\mathbf{a}) - e_0(y_u|y_b). \tag{A69}$$

where $e_0(y|y_b)$ is defined in equation (A75). Combining equations (A75) and (A69), we can therefore express the equity claim as:

$$e(y_t|\mathbf{a}) \quad = \quad e_0(y_t|y_b) + p_u(y_t|\mathbf{a})\,\Gamma(\mathbf{a}). \tag{A70}$$

Because $p_u(y_t|\mathbf{a})$ is positive for all values of $y_t \in (y_b, y_u)$, equation (A70) shows that if $\Gamma(\mathbf{a})$ is positive, then the equity value for a firm that follows the policy $\mathbf{a} = (y_b, y_u, \hat{y})$ is higher than the equity value for a firm that follows a static capital structure policy for *all* values of $y_t \in (y_b, y_u)$. Differentiating equation (A70) we find:

$$e'(y_b|\mathbf{a}) = (1-\tau)(1 - K(y_0, y_b)) + \left(\frac{\Theta - \Pi}{y_b}\right)\left(\frac{\Gamma(\mathbf{a})}{\left(\frac{y_u}{y_b}\right)^{\Theta} - \left(\frac{y_u}{y_b}\right)^{\Pi}}\right). \tag{A71}$$

where

$$K(y_0, y_b) \equiv \left(\frac{y_0}{y_b}\right)(1 - \Pi) + \Pi. \tag{A72}$$

If y_b satisfies the smooth pasting condition, $e'(y_b|\mathbf{a}) = 0$ and from (A71) we have

$$\left.\frac{\Gamma(\mathbf{a})}{\left(\frac{y_u}{y_b}\right)^{\Theta} - \left(\frac{y_u}{y_b}\right)^{\Pi}}\right|_{\mathrm{SP}} = \frac{(1-\tau)(K(y_0, y_b) - 1)y_b}{\Theta - \Pi}. \tag{A73}$$

Then, combining equations (A73)), (A75), (A70), and (A73)), we find that the equity value $e_{\mathrm{SP}}(y_t|\mathbf{a})$ corresponding to a barrier strategy in which y_b is chosen to satisfy the smooth-pasting condition $e'(y_b) = 0$ is

$$\begin{aligned} e_{\mathrm{SP}}(y_t|\mathbf{a}) &= (1-\tau)\left[\frac{y_t}{r-\mu} - \frac{y_0}{r-\mu}\left(\frac{\Pi-1}{\Pi}\right) + \frac{(K(y_0,y_b)-1)y_b}{(\Theta-\Pi)(r-\mu)}\left(\frac{y_t}{y_b}\right)^{\Theta}\right. \\ &\quad \left. + \left(\left(\frac{1}{-\Pi}\right)\frac{y_b}{r-\mu}K(y_0,y_b) - \frac{(K(y_0,y_b)-1)y_b}{\Theta-\Pi}\right)\left(\frac{y_t}{y_b}\right)^{\Pi}\right]. \end{aligned} \tag{A74}$$

This implies that, for an exogenously specified $(\hat{y}, y_u)$, when the boundary y_b satisfies the smooth-pasting condition, the value of equity $e_{\mathrm{SP}}(y_t|\mathbf{a})$ is independent of both $(y_u, \hat{y})$. Twice differentiating equation (A74), and assuming $K(y_0, y_b) > 1$ (i.e., $y_b < y_0$) we find, that for all y_t,

$$\begin{aligned} e'_{\mathrm{SP}}(y_t|\mathbf{a}) &= (1-\tau)\left\{\left(1-\left(\frac{y_t}{y_b}\right)^{\Pi-1}\right) + \left(\frac{\Theta(K(y_0,y_b)-1)}{\Theta-\Pi}\right)\left[\left(\frac{y_t}{y_b}\right)^{\Theta-1} - \left(\frac{y_t}{y_b}\right)^{\Pi-1}\right]\right\} > 0, \\ e''_{\mathrm{SP}}(y_t|\mathbf{a}) &= (1-\tau)\left\{\left(\frac{\Theta}{\Theta-\Pi}\right)\left(\frac{K(y_0,y_b)-1}{y_b}\right)(\Theta-1)\left(\frac{y_t}{y_b}\right)^{\Theta-2}\right. \\ &\quad \left. + \left(\frac{1-\Pi}{y_b}\right)\left(\frac{\Theta K(y_0,y_b)-\Pi}{\Theta-\Pi}\right)\left(\frac{y_t}{y_b}\right)^{\Pi-2}\right\} > 0, \end{aligned}$$

implying that the the equity function is increasing and convex for all y. □

A.1 The HJB equation

To formally characterize the MPE in the presence of debt issuance we rely on a generalization of the standard Hamilton-Jacobi-Bellman equation to handle possibly non-differentiable functions. In this general setting we will be looking for solutions of the HJB equation *in the distributional sense* or *weak solutions* (Crandall and Lions (1983)).

If $v : \mathbb{R}_+ \to \mathbb{R}$ is a convex function then its one sided derivatives $v'_\pm(y)$ are nondecreasing functions of finite variation, and its second distributional derivative is a positive measure that we denote by $v''(dy)$. Consider now the measure

$$\mathcal{O}v(dy) = [(\xi+\mu)yv'_-(y) - \rho v(y)]dy + \frac{1}{2}\sigma^2 y^2 v''(dy). \tag{A75}$$

Lamberton, Zervos, et al. (2013) show that the solution to (A20) or, equivalently, (A24) is intimately related to the set of functions that solve the HJB equation

$$\max\{\mathcal{O}v(y), \psi(y|\mathbf{a}) - v(y)\} = 0 \tag{A76}$$

in the distributional sense. To make this result precise we start by formally defining the type of weak solutions we are interested in.

Definition 5. *A function $v : (0,\infty) \to \mathbb{R}$ is a solution to* (A76) *in the sense of distributions if it is convex and such that*

i) $v(y) \geq \psi(y|\mathbf{a})$ *for all* $y \geq 0$

ii) $\mathcal{O}v(dy)$ *is a non positive measure on* $\mathbb{R}_+$

iii) $\mathcal{O}v(dy)$ *does not charge the set* $\{y \geq 0 : v(y) > \psi(y|\mathbf{a})\}$

Proposition 11 (HJB characterization)**.** $\mathbf{a} \in \mathcal{M}_r \cap \mathcal{S}$ *is a* rMPE *if and only if*

$$v(y) \equiv e(y|\mathbf{a}) - \hat{e}(y) \tag{A77}$$

solves (A76) *in the sense of distributions subject to the boundary conditions*

$$\limsup_{y\downarrow 0} y^{-\Pi} v(y) = \limsup_{y\downarrow 0} y^{-\Pi} \psi(y|\mathbf{a}) < \infty, \tag{A78}$$

$$\limsup_{y\uparrow\infty} y^{-\Theta} v(y) = \limsup_{y\uparrow\infty} y^{-\Theta} \psi(y|\mathbf{a}) < \infty, \tag{A79}$$

where $\Pi < 0$ *and* $\Theta > 1$ *are given in equations* (3.15) *and* (3.14).

Proof This follows from Proposition 8 and Lamberton, Zervos, et al. (2013, Theorems 6.3|4) using the fact that in our case the state space is the positive real line with inaccessible boundaries and the reward function is convex and thus continuous. $\square$

The next proposition provides a characterization of the equity value in a MPE in barrier strategies as a weak solution of the HJB equation.

Proposition 12. *A barrier strategy is a* rMPE *if and only if the induced equity value function* $e(y|\mathbf{a})$ *is a solution to* (A76) *in the sense of distributions.*

Proof. Combining Corollary 2, and the characterization of bond and equity values in equations (3.13) and (3.27), respectively, we deduce that there exists a constant $k > 0$ such that

$$|e(y|\mathbf{a})| \vee |\phi(y|\mathbf{a})| \leq k\,(1 + |y|)\,, \qquad y \geq 0. \tag{A80}$$

Since $\Pi < 0$ and $\Theta > 1$ this implies that we have

$$\lim_{y\downarrow 0} y^{-\Pi} f(y) = \lim_{y\uparrow\infty} y^{-\Theta} f(y) = 0, \qquad \text{for } f \in \{e(\cdot|\mathbf{a}), \phi(\cdot|\mathbf{a})^+\}. \tag{A81}$$

This shows that the boundary conditions (A78) and (A79) hold for any barrier strategy and the desired result now follows from Proposition 11. $\square$